(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
25 April 2002 (25.04.2002)

PCT

(10) International Publication Number
WO 02/33086 A2

(51) International Patent Classification[7]: C12N 15/12, C07K 14/47, 14/705, 16/18, 16/28, C12Q 1/68, G01N 33/68, C12N 15/63, 1/21, 5/10, A01K 67/027

(21) International Application Number: PCT/US01/32152

(22) International Filing Date: 17 October 2001 (17.10.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/240,836	17 October 2000 (17.10.2000)	US
09/804,474	13 March 2001 (13.03.2001)	US

(71) Applicant: **PE CORPORATION (NY)** [US/US]; Global Headquarters, 301 Merritt 7, P.O. Box 5435, Norwalk, CT 06856-5435 (US).

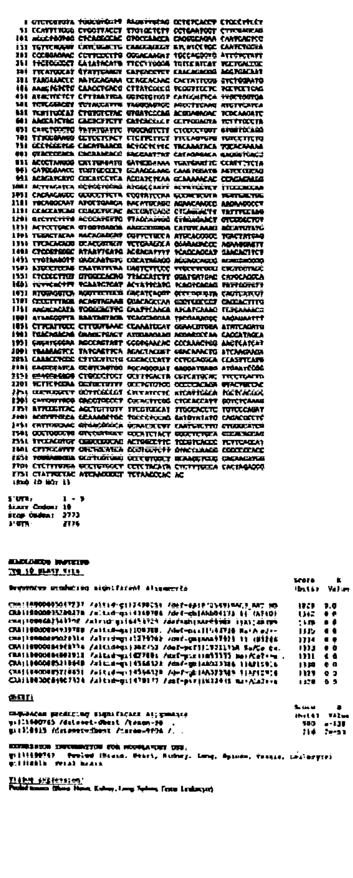
(72) Inventors: **MERKULOV, Gennady, V.**; c/o Celera Genomics, 45 West Gude Drive C2-4#21, Rockville, MD 20850 (US). **KETCHUM, Karen, A.**; c/o Celera Genomics, 45 West Gude Drive C2-4#21, Rockville, MD 20850 (US). **SHAO, Wei**; c/o Celera Genomics, 45 West Gude Drive C2-4#21, Rockville, MD 20850 (US). **YAN, Chunhua**; c/o Celera Genomics, 45 West Gude Drive C2-4#21, Rockville, MD 20850 (US). **DI FRANCESCO, Valentina**; c/o Celera Genomics, 45 West Gude Drive C2-4#21, Rockville, MD 20850 (US). **BEASLEY, Ellen, M.**; c/o Celera Genomics, 45 West Gude Drive C2-4#21, Rockville, MD 20850 (US).

(74) Agent: **MILLMAN, Robert, A.**; Celera Genomics, 45 West Gude Drive C2-4#21, Rockville, MD 20850 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,

[Continued on next page]

(54) Title: ISOLATED HUMAN TRANSPORTER PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS, AND USES THEREOF



(57) Abstract: The present invention provides amino acid sequences of peptides that are encoded by genes within the human genome, the transporter peptides of the present invention. The present invention specifically provides isolated peptide and nucleic acid molecules, methods of identifying orthologs and paralogs of the transporter peptides, and methods of identifying modulators of the transporter peptides.

WO 02/33086 A2

GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

— *entirely in electronic form (except for this front page) and available upon request from the International Bureau*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

ISOLATED HUMAN TRANSPORTER PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS, AND USES THEREOF

RELATED APPLICATIONS

The present application claims priority to provisional application U.S. Serial No. 60/240,836, filed October 17, 2000 (Atty. Docket CL000891-PROV) and 09/804,474, filed March 13, 2001(Atty. Docket CL000891).

FIELD OF THE INVENTION

The present invention is in the field of transporter proteins that are related to the sodium/calcium exchanger subfamily, recombinant DNA molecules, and protein production. The present invention specifically provides novel peptides and proteins that effect ligand transport and nucleic acid molecules encoding such peptide and protein molecules, all of which are useful in the development of human therapeutics and diagnostic compositions and methods.

BACKGROUND OF THE INVENTION

Transporters

Transporter proteins regulate many different functions of a cell, including cell proliferation, differentiation, and signaling processes, by regulating the flow of molecules such as ions and macromolecules, into and out of cells. Transporters are found in the plasma membranes of virtually every cell in eukaryotic organisms. Transporters mediate a variety of cellular functions including regulation of membrane potentials and absorption and secretion of molecules and ion across cell membranes. When present in intracellular membranes of the Golgi apparatus and endocytic vesicles, transporters, such as chloride channels, also regulate organelle pH. For a review, see Greger, R. (1988) Annu. Rev. Physiol. 50:111-122.

Transporters are generally classified by structure and the type of mode of action. In addition, transporters are sometimes classified by the molecule type that is transported, for example, sugar transporters, chlorine channels, potassium channels, etc. There may be many classes of channels for transporting a single type of molecule (a detailed review of channel types can be found at Alexander, S.P.H. and J.A. Peters: Receptor and transporter nomenclature

supplement. Trends Pharmacol. Sci., Elsevier, pp. 65-68 (1997) and http://www-biology.ucsd.edu/~msaier/transport/titlepage2.html.

The following general classification scheme is known in the art and is followed in the present discoveries.

Channel-type transporters. Transmembrane channel proteins of this class are ubiquitously found in the membranes of all types of organisms from bacteria to higher eukaryotes. Transport systems of this type catalyze facilitated diffusion (by an energy-independent process) by passage through a transmembrane aqueous pore or channel without evidence for a carrier-mediated mechanism. These channel proteins usually consist largely of a-helical spanners, although b-strands may also be present and may even comprise the channel. However, outer membrane porin-type channel proteins are excluded from this class and are instead included in class 9.

Carrier-type transporters. Transport systems are included in this class if they utilize a carrier-mediated process to catalyze uniport (a single species is transported by facilitated diffusion), antiport (two or more species are transported in opposite directions in a tightly coupled process, not coupled to a direct form of energy other than chemiosmotic energy) and/or symport (two or more species are transported together in the same direction in a tightly coupled process, not coupled to a direct form of energy other than chemiosmotic energy).

Pyrophosphate bond hydrolysis-driven active transporters. Transport systems are included in this class if they hydrolyze pyrophosphate or the terminal pyrophosphate bond in ATP or another nucleoside triphosphate to drive the active uptake and/or extrusion of a solute or solutes. The transport protein may or may not be transiently phosphorylated, but the substrate is not phosphorylated.

PEP-dependent, phosphoryl transfer-driven group translocators. Transport systems of the bacterial phosphoenolpyruvate:sugar phosphotransferase system are included in this class. The product of the reaction, derived from extracellular sugar, is a cytoplasmic sugar-phosphate.

Decarboxylation-driven active transporters. Transport systems that drive solute (e.g., ion) uptake or extrusion by decarboxylation of a cytoplasmic substrate are included in this class.

Oxidoreduction-driven active transporters. Transport systems that drive transport of a solute (e.g., an ion) energized by the flow of electrons from a reduced substrate to an oxidized substrate are included in this class.

Light-driven active transporters. Transport systems that utilize light energy to drive transport of a solute (e.g., an ion) are included in this class.

Mechanically-driven active transporters. Transport systems are included in this class if they drive movement of a cell or organelle by allowing the flow of ions (or other solutes) through the membrane down their electrochemical gradients.

Outer-membrane porins (of b-structure). These proteins form transmembrane pores or channels that usually allow the energy independent passage of solutes across a membrane. The transmembrane portions of these proteins consist exclusively of b-strands that form a b-barrel. These porin-type proteins are found in the outer membranes of Gram-negative bacteria, mitochondria and eukaryotic plastids.

Methyltransferase-driven active transporters. A single characterized protein currently falls into this category, the Na+-transporting methyltetrahydromethanopterin:coenzyme M methyltransferase.

Non-ribosome-synthesized channel-forming peptides or peptide-like molecules. These molecules, usually chains of L- and D-amino acids as well as other small molecular building blocks such as lactate, form oligomeric transmembrane ion channels. Voltage may induce channel formation by promoting assembly of the transmembrane channel. These peptides are often made by bacteria and fungi as agents of biological warfare.

Non-Proteinaceous Transport Complexes. Ion conducting substances in biological membranes that do not consist of or are not derived from proteins or peptides fall into this category.

Functionally characterized transporters for which sequence data are lacking. Transporters of particular physiological significance will be included in this category even though a family assignment cannot be made.

Putative transporters in which no family member is an established transporter. Putative transport protein families are grouped under this number and will either be classified elsewhere when the transport function of a member becomes established, or will be eliminated from the TC classification system if the proposed transport function is disproven. These families include a member or members for which a transport function has been suggested, but evidence for such a function is not yet compelling.

Auxiliary transport proteins. Proteins that in some way facilitate transport across one or more biological membranes but do not themselves participate directly in transport are included in this class. These proteins always function in conjunction with one or more transport proteins. They may provide a function connected with energy coupling to transport, play a structural role in complex formation or serve a regulatory function.

Transporters of unknown classification. Transport protein families of unknown classification are grouped under this number and will be classified elsewhere when the transport process and energy coupling mechanism are characterized. These families include at least one member for which a transport function has been established, but either the mode of transport or the energy coupling mechanism is not known.

Ion channels

An important type of transporter is the ion channel. Ion channels regulate many different cell proliferation, differentiation, and signaling processes by regulating the flow of ions into and out of cells. Ion channels are found in the plasma membranes of virtually every cell in eukaryotic organisms. Ion channels mediate a variety of cellular functions including regulation of membrane potentials and absorption and secretion of ion across epithelial membranes. When present in intracellular membranes of the Golgi apparatus and endocytic vesicles, ion channels, such as chloride channels, also regulate organelle pH. For a review, see Greger, R. (1988) Annu. Rev. Physiol. 50:111-122.

Ion channels are generally classified by structure and the type of mode of action. For example, extracellular ligand gated channels (ELGs) are comprised of five polypeptide subunits, with each subunit having 4 membrane spanning domains, and are activated by the binding of an extracellular ligand to the channel. In addition, channels are sometimes classified by the ion type that is transported, for example, chlorine channels, potassium channels, etc. There may be many classes of channels for transporting a single type of ion (a detailed review of channel types can be found at Alexander, S.P.H. and J.A. Peters (1997). Receptor and ion channel nomenclature supplement. Trends Pharmacol. Sci., Elsevier, pp. 65-68 and http://www-biology.ucsd.edu/~msaier/transport/toc.html.

There are many types of ion channels based on structure. For example, many ion channels fall within one of the following groups: extracellular ligand-gated channels (ELG), intracellular ligand-gated channels (ILG), inward rectifying channels (INR), intercellular (gap junction) channels, and voltage gated channels (VIC). There are additionally recognized other channel families based on ion-type transported, cellular location and drug sensitivity. Detailed information on each of these, their activity, ligand type, ion type, disease association, drugability, and other information pertinent to the present invention, is well known in the art.

Extracellular ligand-gated channels, ELGs, are generally comprised of five polypeptide subunits, Unwin, N. (1993), Cell 72: 31-41; Unwin, N. (1995), Nature 373: 37-43; Hucho, F., et

al., (1996) J. Neurochem. 66: 1781-1792; Hucho, F., et al., (1996) Eur. J. Biochem. 239: 539-557; Alexander, S.P.H. and J.A. Peters (1997), Trends Pharmacol. Sci., Elsevier, pp. 4-6; 36-40; 42-44; and Xue, H. (1998) J. Mol. Evol. 47: 323-333. Each subunit has 4 membrane spanning regions: this serves as a means of identifying other members of the ELG family of proteins. ELG bind a ligand and in response modulate the flow of ions. Examples of ELG include most members of the neurotransmitter-receptor family of proteins, e.g., GABAI receptors. Other members of this family of ion channels include glycine receptors, ryandyne receptors, and ligand gated calcium channels.

Sodium/Calcium Exchangers

The protein provided by the present invention is a novel sodium/calcium exchanger. Sodium/calcium exchangers (NCX) rapidly import calcium during excitation impulse. Intracellular calcium concentrations vary greatly during the excitation/relaxation cycle. In contrast, extracellular calcium concentrations are maintained at relatively steady levels, despite wide variations in the amounts of calcium supplied with food.

There are at least three known mammalian NCX genes and a number of alternatively spliced isoforms. NCX sequences are highly conserved. NCX proteins contain 9 transmembrane domains and are regulated by calcium and sodium ions and, to some extent, by phosphorylation.

NCX proteins initiate cardiac myocyte contractions; this effect has been confirmed by *in vitro* experiments. Together with calsequestrin, a calcium binding protein, NCX proteins maintain calcium homeostasis in the heart muscle. This regulatory mechanism depends on the gene dosage, as evident from experiments with transgenic animals. Variations in expression levels of these proteins may be associated with some forms of heart disease.

Calcium transporters can mediate divalent ion toxicity. Barium and strontium can be carried by these channels into the cell, albeit at slower rates than calcium, which is the natural substrate. A panel of bivalent cations, such as copper, lead, cadmium, cobalt and nickel, inhibit calcium flow, but do not penetrate the cell membrane. Bivalent and trivalent iron, manganese, and zinc show no effect.

The sequence of the sodium/calcium exchanger provided by the present invention may be used to screen human populations for mutations associated with neurological conditions and heart disease. Furthermore, drugs can be designed that target this and other transporters.

For a further review of sodium/calcium exchangers, see: Linck *et al.*, *J Pharmacol Exp Ther* 2000 Aug;294(2):648-57; Shen *et al.*, *J Pharmacol Exp Ther* 2000 Aug;294(2):562-70;

Philipson *et al., Annu Rev Physiol* 2000;62:111-33; Zhang *et al., Br J Pharmacol* 2000 Jun;130(3):485-8; and Vercesi *et al., FEBS Lett* 2000 May 12;473(2):203-6.

The Voltage-gated Ion Channel (VIC) Superfamily

Proteins of the VIC family are ion-selective channel proteins found in a wide range of bacteria, archaea and eukaryotes Hille, B. (1992), Chapter 9: Structure of channel proteins; Chapter 20: Evolution and diversity. In: Ionic Channels of Excitable Membranes, 2nd Ed., Sinaur Assoc. Inc., Pubs., Sunderland, Massachusetts; Sigworth, F.J. (1993), Quart. Rev. Biophys. 27: 1-40; Salkoff, L. and T. Jegla (1995), Neuron 15: 489-492; Alexander, S.P.H. et al., (1997), Trends Pharmacol. Sci., Elsevier, pp. 76-84; Jan, L.Y. et al., (1997), Annu. Rev. Neurosci. 20: 91-123; Doyle, D.A, et al., (1998) Science 280: 69-77; Terlau, H. and W. Stühmer (1998), Naturwissenschaften 85: 437-444. They are often homo- or heterooligomeric structures with several dissimilar subunits (e.g., a1-a2-d-b Ca^{2+} channels, ab_1b_2 Na^+ channels or $(a)_4$-b K^+ channels), but the channel and the primary receptor is usually associated with the a (or a1) subunit. Functionally characterized members are specific for K^+, Na^+ or Ca^{2+}. The K^+ channels usually consist of homotetrameric structures with each a-subunit possessing six transmembrane spanners (TMSs). The a1 and a subunits of the Ca^{2+} and Na^+ channels, respectively, are about four times as large and possess 4 units, each with 6 TMSs separated by a hydrophilic loop, for a total of 24 TMSs. These large channel proteins form heterotetra-unit structures equivalent to the homotetrameric structures of most K^+ channels. All four units of the Ca^{2+} and Na^+ channels are homologous to the single unit in the homotetrameric K^+ channels. Ion flux via the eukaryotic channels is generally controlled by the transmembrane electrical potential (hence the designation, voltage-sensitive) although some are controlled by ligand or receptor binding.

Several putative K^+-selective channel proteins of the VIC family have been identified in prokaryotes. The structure of one of them, the KcsA K^+ channel of *Streptomyces lividans*, has been solved to 3.2 Å resolution. The protein possesses four identical subunits, each with two transmembrane helices, arranged in the shape of an inverted teepee or cone. The cone cradles the "selectivity filter" P domain in its outer end. The narrow selectivity filter is only 12 Å long, whereas the remainder of the channel is wider and lined with hydrophobic residues. A large water-filled cavity and helix dipoles stabilize K^+ in the pore. The selectivity filter has two bound K^+ ions about 7.5 Å apart from each other. Ion conduction is proposed to result from a balance of electrostatic attractive and repulsive forces.

In eukaryotes, each VIC family channel type has several subtypes based on pharmacological and electrophysiological data. Thus, there are five types of Ca^{2+} channels (L, N, P, Q and T). There are at least ten types of K^+ channels, each responding in different ways to different stimuli: voltage-sensitive [Ka, Kv, Kvr, Kvs and Ksr], Ca^{2+}-sensitive [BK_{Ca}, IK_{Ca} and SK_{Ca}] and receptor-coupled [K_M and K_{ACh}]. There are at least six types of Na^+ channels (I, II, III, μ1, H1 and PN3). Tetrameric channels from both prokaryotic and eukaryotic organisms are known in which each a-subunit possesses 2 TMSs rather than 6, and these two TMSs are homologous to TMSs 5 and 6 of the six TMS unit found in the voltage-sensitive channel proteins. KcsA of *S. lividans* is an example of such a 2 TMS channel protein. These channels may include the K_{Na} (Na^+-activated) and K_{Vol} (cell volume-sensitive) K^+ channels, as well as distantly related channels such as the Tok1 K^+ channel of yeast, the TWIK-1 inward rectifier K^+ channel of the mouse and the TREK-1 K^+ channel of the mouse. Because of insufficient sequence similarity with proteins of the VIC family, inward rectifier K^+ IRK channels (ATP-regulated; G-protein-activated) which possess a P domain and two flanking TMSs are placed in a distinct family. However, substantial sequence similarity in the P region suggests that they are homologous. The b, g and d subunits of VIC family members, when present, frequently play regulatory roles in channel activation/deactivation.

<u>The Epithelial Na^+ Channel (ENaC) Family</u>

The ENaC family consists of over twenty-four sequenced proteins (Canessa, C.M., et al., (1994), Nature 367: 463-467, Le, T. and M.H. Saier, Jr. (1996), Mol. Membr. Biol. 13: 149-157; Garty, H. and L.G. Palmer (1997), Physiol. Rev. 77: 359-396; Waldmann, R., et al., (1997), Nature 386: 173-177; Darboux, I., et al., (1998), J. Biol. Chem. 273: 9424-9429; Firsov, D., et al., (1998), EMBO J. 17: 344-352; Horisberger, J.-D. (1998). Curr. Opin. Struc. Biol. 10: 443-449). All are from animals with no recognizable homologues in other eukaryotes or bacteria. The vertebrate ENaC proteins from epithelial cells cluster tightly together on the phylogenetic tree: voltage-insensitive ENaC homologues are also found in the brain. Eleven sequenced *C. elegans* proteins, including the degenerins, are distantly related to the vertebrate proteins as well as to each other. At least some of these proteins form part of a mechano-transducing complex for touch sensitivity. The homologous *Helix aspersa* (FMRF-amide)-activated Na^+ channel is the first peptide neurotransmitter-gated ionotropic receptor to be sequenced.

Protein members of this family all exhibit the same apparent topology, each with N- and C-termini on the inside of the cell, two amphipathic transmembrane spanning segments, and a

large extracellular loop. The extracellular domains contain numerous highly conserved cysteine residues. They are proposed to serve a receptor function.

Mammalian ENaC is important for the maintenance of Na^+ balance and the regulation of blood pressure. Three homologous ENaC subunits, alpha, beta, and gamma, have been shown to assemble to form the highly Na^+-selective channel. The stoichiometry of the three subunits is $alpha_2$, beta1, gamma1 in a heterotetrameric architecture.

The Glutamate-gated Ion Channel (GIC) Family of Neurotransmitter Receptors

Members of the GIC family are heteropentameric complexes in which each of the 5 subunits is of 800-1000 amino acyl residues in length (Nakanishi, N., et al, (1990), Neuron 5: 569-581; Unwin, N. (1993), Cell 72: 31-41; Alexander, S.P.H. and J.A. Peters (1997) Trends Pharmacol. Sci., Elsevier, pp. 36-40). These subunits may span the membrane three or five times as putative a-helices with the N-termini (the glutamate-binding domains) localized extracellularly and the C-termini localized cytoplasmically. They may be distantly related to the ligand-gated ion channels, and if so, they may possess substantial b-structure in their transmembrane regions. However, homology between these two families cannot be established on the basis of sequence comparisons alone. The subunits fall into six subfamilies: a, b, g, d, e and z.

The GIC channels are divided into three types: (1) a-amino-3-hydroxy-5-methyl-4-isoxazole propionate (AMPA)-, (2) kainate- and (3) N-methyl-D-aspartate (NMDA)-selective glutamate receptors. Subunits of the AMPA and kainate classes exhibit 35-40% identity with each other while subunits of the NMDA receptors exhibit 22-24% identity with the former subunits. They possess large N-terminal, extracellular glutamate-binding domains that are homologous to the periplasmic glutamine and glutamate receptors of ABC-type uptake permeases of Gram-negative bacteria. All known members of the GIC family are from animals. The different channel (receptor) types exhibit distinct ion selectivities and conductance properties. The NMDA-selective large conductance channels are highly permeable to monovalent cations and Ca^{2+}. The AMPA- and kainate-selective ion channels are permeable primarily to monovalent cations with only low permeability to Ca^{2+}.

The Chloride Channel (ClC) Family

The ClC family is a large family consisting of dozens of sequenced proteins derived from Gram-negative and Gram-positive bacteria, cyanobacteria, archaea, yeast, plants and animals (Steinmeyer, K., et al., (1991), Nature 354: 301-304; Uchida, S., et al., (1993), J. Biol. Chem.

268: 3821-3824; Huang, M.-E., et al., (1994), J. Mol. Biol. 242: 595-598; Kawasaki, M., et al, (1994), Neuron 12: 597-604; Fisher, W.E., et al., (1995), Genomics. 29:598-606; and Foskett, J.K. (1998), Annu. Rev. Physiol. 60: 689-717). These proteins are essentially ubiquitous, although they are not encoded within genomes of *Haemophilus influenzae*, *Mycoplasma genitalium*, and *Mycoplasma pneumoniae*. Sequenced proteins vary in size from 395 amino acyl residues *(M. jannaschii)* to 988 residues (man). Several organisms contain multiple ClC family paralogues. For example, *Synechocystis* has two paralogues, one of 451 residues in length and the other of 899 residues. *Arabidopsis thaliana* has at least four sequenced paralogues, (775-792 residues), humans also have at least five paralogues (820-988 residues), and *C. elegans* also has at least five (810-950 residues). There are nine known members in mammals, and mutations in three of the corresponding genes cause human diseases. *E. coli*, *Methanococcus jannaschii* and *Saccharomyces cerevisiae* only have one ClC family member each. With the exception of the larger *Synechocystis* paralogue, all bacterial proteins are small (395-492 residues) while all eukaryotic proteins are larger (687-988 residues). These proteins exhibit 10-12 putative transmembrane a-helical spanners (TMSs) and appear to be present in the membrane as homodimers. While one member of the family, *Torpedo* ClC-O, has been reported to have two channels, one per subunit, others are believed to have just one.

All functionally characterized members of the ClC family transport chloride, some in a voltage-regulated process. These channels serve a variety of physiological functions (cell volume regulation; membrane potential stabilization; signal transduction; transepithelial transport, etc.). Different homologues in humans exhibit differing anion selectivities, i.e., ClC4 and ClC5 share a $NO_3^- > Cl^- > Br^- > I^-$ conductance sequence, while ClC3 has an $I^- > Cl^-$ selectivity. The ClC4 and ClC5 channels and others exhibit outward rectifying currents with currents only at voltages more positive than +20mV.

Animal Inward Rectifier K^+ Channel (IRK-C) Family

IRK channels possess the "minimal channel-forming structure" with only a P domain, characteristic of the channel proteins of the VIC family, and two flanking transmembrane spanners (Shuck, M.E., et al., (1994), J. Biol. Chem. 269: 24261-24270; Ashen, M.D., et al., (1995), Am. J. Physiol. 268: H506-H511; Salkoff, L. and T. Jegla (1995), Neuron 15: 489-492; Aguilar-Bryan, L., et al., (1998), Physiol. Rev. 78: 227-245; Ruknudin, A., et al., (1998), J. Biol. Chem. 273: 14165-14171). They may exist in the membrane as homo- or heterooligomers. They have a greater tendency to let K^+ flow into the cell than out. Voltage-dependence may be regulated by external K^+, by internal Mg^{2+}, by internal ATP and/or by G-proteins. The P domains

of IRK channels exhibit limited sequence similarity to those of the VIC family, but this sequence similarity is insufficient to establish homology. Inward rectifiers play a role in setting cellular membrane potentials, and the closing of these channels upon depolarization permits the occurrence of long duration action potentials with a plateau phase. Inward rectifiers lack the intrinsic voltage sensing helices found in VIC family channels. In a few cases, those of Kir1.1a and Kir6.2, for example, direct interaction with a member of the ABC superfamily has been proposed to confer unique functional and regulatory properties to the heteromeric complex, including sensitivity to ATP. The SUR1 sulfonylurea receptor (spQ09428) is the ABC protein that regulates the Kir6.2 channel in response to ATP, and CFTR may regulate Kir1.1a. Mutations in SUR1 are the cause of familial persistent hyperinsulinemic hypoglycemia in infancy (PHHI), an autosomal recessive disorder characterized by unregulated insulin secretion in the pancreas.

ATP-gated Cation Channel (ACC) Family

Members of the ACC family (also called P2X receptors) respond to ATP, a functional neurotransmitter released by exocytosis from many types of neurons (North, R.A. (1996), Curr. Opin. Cell Biol. 8: 474-483; Soto, F., M. Garcia-Guzman and W. Stühmer (1997), J. Membr. Biol. 160: 91-100). They have been placed into seven groups ($P2X_1$ - $P2X_7$) based on their pharmacological properties. These channels, which function at neuron-neuron and neuron-smooth muscle junctions, may play roles in the control of blood pressure and pain sensation. They may also function in lymphocyte and platelet physiology. They are found only in animals.

The proteins of the ACC family are quite similar in sequence (>35% identity), but they possess 380-1000 amino acyl residues per subunit with variability in length localized primarily to the C-terminal domains. They possess two transmembrane spanners, one about 30-50 residues from their N-termini, the other near residues 320-340. The extracellular receptor domains between these two spanners (of about 270 residues) are well conserved with numerous conserved glycyl and cysteyl residues. The hydrophilic C-termini vary in length from 25 to 240 residues. They resemble the topologically similar epithelial Na^+ channel (ENaC) proteins in possessing (a) N- and C-termini localized intracellularly, (b) two putative transmembrane spanners, (c) a large extracellular loop domain, and (d) many conserved extracellular cysteyl residues. ACC family members are, however, not demonstrably homologous with them. ACC channels are probably hetero- or homomultimers and transport small monovalent cations (Me^+). Some also transport Ca^{2+}; a few also transport small metabolites.

<u>The Ryanodine-Inositol 1,4,5-triphosphate Receptor Ca^{2+} Channel (RIR-CaC) Family</u>

Ryanodine (Ry)-sensitive and inositol 1,4,5-triphosphate (IP3)-sensitive Ca^{2+}-release channels function in the release of Ca^{2+} from intracellular storage sites in animal cells and thereby regulate various Ca^{2+} -dependent physiological processes (Hasan, G. et al., (1992) Development 116: 967-975; Michikawa, T., et al., (1994), J. Biol. Chem. 269: 9184-9189; Tunwell, R.E.A., (1996), Biochem. J. 318: 477-487; Lee, A.G. (1996) *Biomembranes,* Vol. 6, Transmembrane Receptors and Channels (A.G. Lee, ed.), JAI Press, Denver, CO., pp 291-326; Mikoshiba, K., et al., (1996) J. Biochem. Biomem. 6: 273-289). Ry receptors occur primarily in muscle cell sarcoplasmic reticular (SR) membranes, and IP3 receptors occur primarily in brain cell endoplasmic reticular (ER) membranes where they effect release of Ca^{2+} into the cytoplasm upon activation (opening) of the channel.

The Ry receptors are activated as a result of the activity of dihydropyridine-sensitive Ca^{2+} channels. The latter are members of the voltage-sensitive ion channel (VIC) family. Dihydropyridine-sensitive channels are present in the T-tubular systems of muscle tissues.

Ry receptors are homotetrameric complexes with each subunit exhibiting a molecular size of over 500,000 daltons (about 5,000 amino acyl residues). They possess C-terminal domains with six putative transmembrane a -helical spanners (TMSs). Putative pore-forming sequences occur between the fifth and sixth TMSs as suggested for members of the VIC family. The large N-terminal hydrophilic domains and the small C-terminal hydrophilic domains are localized to the cytoplasm. Low resolution 3-dimensional structural data are available. Mammals possess at least three isoforms that probably arose by gene duplication and divergence before divergence of the mammalian species. Homologues are present in humans and *Caenorabditis elegans.*

IP_3 receptors resemble Ry receptors in many respects. (1) They are homotetrameric complexes with each subunit exhibiting a molecular size of over 300,000 daltons (about 2,700 amino acyl residues). (2) They possess C-terminal channel domains that are homologous to those of the Ry receptors. (3) The channel domains possess six putative TMSs and a putative channel lining region between TMSs 5 and 6. (4) Both the large N-terminal domains and the smaller C-terminal tails face the cytoplasm. (5) They possess covalently linked carbohydrate on extracytoplasmic loops of the channel domains. (6) They have three currently recognized isoforms (types 1, 2, and 3) in mammals which are subject to differential regulation and have different tissue distributions.

IP_3 receptors possess three domains: N-terminal IP_3-binding domains, central coupling or regulatory domains and C-terminal channel domains. Channels are activated by IP_3 binding, and like the Ry receptors, the activities of the IP_3 receptor channels are regulated by phosphorylation of the regulatory domains, catalyzed by various protein kinases. They predominate in the endoplasmic reticular membranes of various cell types in the brain but have also been found in the plasma membranes of some nerve cells derived from a variety of tissues.

The channel domains of the Ry and IP_3 receptors comprise a coherent family that in spite of apparent structural similarities, do not show appreciable sequence similarity of the proteins of the VIC family. The Ry receptors and the IP_3 receptors cluster separately on the RIR-CaC family tree. They both have homologues in *Drosophila.* Based on the phylogenetic tree for the family, the family probably evolved in the following sequence: (1) A gene duplication event occurred that gave rise to Ry and IP_3 receptors in invertebrates. (2) Vertebrates evolved from invertebrates. (3) The three isoforms of each receptor arose as a result of two distinct gene duplication events. (4) These isoforms were transmitted to mammals before divergence of the mammalian species.

<u>The Organellar Chloride Channel (O-ClC) Family</u>

Proteins of the O-ClC family are voltage-sensitive chloride channels found in intracellular membranes but not the plasma membranes of animal cells (Landry, D, et al., (1993), J. Biol. Chem. 268: 14948-14955; Valenzuela, Set al., (1997), J. Biol. Chem. 272: 12575-12582; and Duncan, R.R., et al., (1997), J. Biol. Chem. 272: 23880-23886).

They are found in human nuclear membranes, and the bovine protein targets to the microsomes, but not the plasma membrane, when expressed in *Xenopus laevis* oocytes. These proteins are thought to function in the regulation of the membrane potential and in transepithelial ion absorption and secretion in the kidney. They possess two putative transmembrane a-helical spanners (TMSs) with cytoplasmic N- and C-termini and a large luminal loop that may be glycosylated. The bovine protein is 437 amino acyl residues in length and has the two putative TMSs at positions 223-239 and 367-385. The human nuclear protein is much smaller (241 residues). A *C. elegans* homologue is 260 residues long.

Transporter proteins, particularly members of the sodium/calcium exchanger subfamily, are a major target for drug action and development. Accordingly, it is valuable to the field of pharmaceutical development to identify and characterize previously unknown transport proteins. The present invention advances the state of the art by providing previously unidentified human transport proteins.

SUMMARY OF THE INVENTION

The present invention is based in part on the identification of amino acid sequences of human transporter peptides and proteins that are related to the sodium/calcium exchanger subfamily, as well as allelic variants and other mammalian orthologs thereof. These unique peptide sequences, and nucleic acid sequences that encode these peptides, can be used as models for the development of human therapeutic targets, aid in the identification of therapeutic proteins, and serve as targets for the development of human therapeutic agents that modulate transporter activity in cells and tissues that express the transporter. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain.

DESCRIPTION OF THE FIGURE SHEETS

FIGURE 1 provides the nucleotide sequence of a cDNA molecule or transcript sequence that encodes the transporter protein of the present invention (SEQ ID NO:1). In addition structure and functional information is provided, such as ATG start, stop and tissue distribution, where available, that allows one to readily determine specific uses of inventions based on this molecular sequence. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain.

FIGURE 2 provides the predicted amino acid sequence of the transporter of the present invention. (SEQ ID NO:2) In addition structure and functional information such as protein family, function, and modification sites is provided where available, allowing one to readily determine specific uses of inventions based on this molecular sequence.

FIGURE 3 provides genomic sequences that span the gene encoding the transporter protein of the present invention (SEQ ID NO: 3). In addition structure and functional information, such as intron/exon structure, promoter location, etc., is provided where available, allowing one to readily determine specific uses of inventions based on this molecular sequence. 140 SNPs, including 6 indels, have been identified in the gene encoding the transporter protein provided by the present invention and are given in Figure 3.

DETAILED DESCRIPTION OF THE INVENTION

General Description

The present invention is based on the sequencing of the human genome. During the sequencing and assembly of the human genome, analysis of the sequence information revealed previously unidentified fragments of the human genome that encode peptides that share structural and/or sequence homology to protein/peptide/domains identified and characterized within the art as being a transporter protein or part of a transporter protein and are related to the sodium/calcium exchanger subfamily. Utilizing these sequences, additional genomic sequences were assembled and transcript and/or cDNA sequences were isolated and characterized. Based on this analysis, the present invention provides amino acid sequences of human transporter peptides and proteins that are related to the sodium/calcium exchanger subfamily, nucleic acid sequences in the form of transcript sequences, cDNA sequences and/or genomic sequences that encode these transporter peptides and proteins, nucleic acid variation (allelic information), tissue distribution of expression, and information about the closest art known protein/peptide/domain that has structural or sequence homology to the transporter of the present invention.

In addition to being previously unknown, the peptides that are provided in the present invention are selected based on their ability to be used for the development of commercially important products and services. Specifically, the present peptides are selected based on homology and/or structural relatedness to known transporter proteins of the sodium/calcium exchanger subfamily and the expression pattern observed . Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain.. The art has clearly established the commercial importance of members of this family of proteins and proteins that have expression patterns similar to that of the present gene. Some of the more specific features of the peptides of the present invention, and the uses thereof, are described herein, particularly in the Background of the Invention and in the annotation provided in the Figures, and/or are known within the art for each of the known sodium/calcium exchanger family or subfamily of transporter proteins.

Specific Embodiments

Peptide Molecules

The present invention provides nucleic acid sequences that encode protein molecules that have been identified as being members of the transporter family of proteins and are related to the sodium/calcium exchanger subfamily (protein sequences are provided in Figure 2, transcript/cDNA sequences are provided in Figures 1 and genomic sequences are provided in Figure 3). The peptide sequences provided in Figure 2, as well as the obvious variants described herein, particularly allelic variants as identified herein and using the information in Figure 3, will be referred herein as the transporter peptides of the present invention, transporter peptides, or peptides/proteins of the present invention.

The present invention provides isolated peptide and protein molecules that consist of, consist essentially of, or comprising the amino acid sequences of the transporter peptides disclosed in the Figure 2, (encoded by the nucleic acid molecule shown in Figure 1, transcript/cDNA or Figure 3, genomic sequence), as well as all obvious variants of these peptides that are within the art to make and use. Some of these variants are described in detail below.

As used herein, a peptide is said to be "isolated" or "purified" when it is substantially free of cellular material or free of chemical precursors or other chemicals. The peptides of the present invention can be purified to homogeneity or other degrees of purity. The level of purification will be based on the intended use. The critical feature is that the preparation allows for the desired function of the peptide, even if in the presence of considerable amounts of other components (the features of an isolated nucleic acid molecule is discussed below).

In some uses, "substantially free of cellular material" includes preparations of the peptide having less than about 30% (by dry weight) other proteins (i.e., contaminating protein), less than about 20% other proteins, less than about 10% other proteins, or less than about 5% other proteins. When the peptide is recombinantly produced, it can also be substantially free of culture medium, i.e., culture medium represents less than about 20% of the volume of the protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of the peptide in which it is separated from chemical precursors or other chemicals that are involved in its synthesis. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of the transporter peptide having less than about 30% (by dry weight) chemical precursors or other chemicals, less than about 20% chemical

precursors or other chemicals, less than about 10% chemical precursors or other chemicals, or less than about 5% chemical precursors or other chemicals.

The isolated transporter peptide can be purified from cells that naturally express it, purified from cells that have been altered to express it (recombinant), or synthesized using known protein synthesis methods. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. For example, a nucleic acid molecule encoding the transporter peptide is cloned into an expression vector, the expression vector introduced into a host cell and the protein expressed in the host cell. The protein can then be isolated from the cells by an appropriate purification scheme using standard protein purification techniques. Many of these techniques are described in detail below.

Accordingly, the present invention provides proteins that consist of the amino acid sequences provided in Figure 2 (SEQ ID NO:2), for example, proteins encoded by the transcript/cDNA nucleic acid sequences shown in Figure 1 (SEQ ID NO:1) and the genomic sequences provided in Figure 3 (SEQ ID NO:3). The amino acid sequence of such a protein is provided in Figure 2. A protein consists of an amino acid sequence when the amino acid sequence is the final amino acid sequence of the protein.

The present invention further provides proteins that consist essentially of the amino acid sequences provided in Figure 2 (SEQ ID NO:2), for example, proteins encoded by the transcript/cDNA nucleic acid sequences shown in Figure 1 (SEQ ID NO:1) and the genomic sequences provided in Figure 3 (SEQ ID NO:3). A protein consists essentially of an amino acid sequence when such an amino acid sequence is present with only a few additional amino acid residues, for example from about 1 to about 100 or so additional residues, typically from 1 to about 20 additional residues in the final protein.

The present invention further provides proteins that comprise the amino acid sequences provided in Figure 2 (SEQ ID NO:2), for example, proteins encoded by the transcript/cDNA nucleic acid sequences shown in Figure 1 (SEQ ID NO:1) and the genomic sequences provided in Figure 3 (SEQ ID NO:3). A protein comprises an amino acid sequence when the amino acid sequence is at least part of the final amino acid sequence of the protein. In such a fashion, the protein can be only the peptide or have additional amino acid molecules, such as amino acid residues (contiguous encoded sequence) that are naturally associated with it or heterologous amino acid residues/peptide sequences. Such a protein can have a few additional amino acid residues or can comprise several hundred or more additional amino acids. The preferred classes of proteins that are comprised of the

transporter peptides of the present invention are the naturally occurring mature proteins. A brief description of how various types of these proteins can be made/isolated is provided below.

The transporter peptides of the present invention can be attached to heterologous sequences to form chimeric or fusion proteins. Such chimeric and fusion proteins comprise a transporter peptide operatively linked to a heterologous protein having an amino acid sequence not substantially homologous to the transporter peptide. "Operatively linked" indicates that the transporter peptide and the heterologous protein are fused in-frame. The heterologous protein can be fused to the N-terminus or C-terminus of the transporter peptide.

In some uses, the fusion protein does not affect the activity of the transporter peptide *per se*. For example, the fusion protein can include, but is not limited to, enzymatic fusion proteins, for example beta-galactosidase fusions, yeast two-hybrid GAL fusions, poly-His fusions, MYC-tagged, HI-tagged and Ig fusions. Such fusion proteins, particularly poly-His fusions, can facilitate the purification of recombinant transporter peptide. In certain host cells (e.g., mammalian host cells), expression and/or secretion of a protein can be increased by using a heterologous signal sequence.

A chimeric or fusion protein can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different protein sequences are ligated together in-frame in accordance with conventional techniques. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and re-amplified to generate a chimeric gene sequence (see Ausubel *et al.*, *Current Protocols in Molecular Biology*, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST protein). A transporter peptide-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the transporter peptide.

As mentioned above, the present invention also provides and enables obvious variants of the amino acid sequence of the proteins of the present invention, such as naturally occurring mature forms of the peptide, allelic/sequence variants of the peptides, non-naturally occurring recombinantly derived variants of the peptides, and orthologs and paralogs of the peptides. Such variants can readily be generated using art-known techniques in the fields of recombinant nucleic acid technology and protein biochemistry. It is understood, however, that variants exclude any amino acid sequences disclosed prior to the invention.

Such variants can readily be identified/made using molecular techniques and the sequence information disclosed herein. Further, such variants can readily be distinguished from other peptides based on sequence and/or structural homology to the transporter peptides of the present invention. The degree of homology/identity present will be based primarily on whether the peptide is a functional variant or non-functional variant, the amount of divergence present in the paralog family and the evolutionary distance between the orthologs.

To determine the percent identity of two amino acid sequences or two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). In a preferred embodiment, at least 30%, 40%, 50%, 60%, 70%, 80%, or 90% or more of a reference sequence is aligned for comparison purposes. The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

The comparison of sequences and determination of percent identity and similarity between two sequences can be accomplished using a mathematical algorithm. (*Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data, Part 1*, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991). In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (*J. Mol. Biol.* (48):444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package (available at http://www.gcg.com), using either a Blossom 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (Devereux, J., *et al., Nucleic Acids Res. 12(1)*:387

(1984)) (available at http://www.gcg.com), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. In another embodiment, the percent identity between two amino acid or nucleotide sequences is determined using the algorithm of E. Myers and W. Miller (CABIOS, 4:11-17 (1989)) which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

The nucleic acid and protein sequences of the present invention can further be used as a "query sequence" to perform a search against sequence databases to, for example, identify other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, *et al.* (*J. Mol. Biol.* 215:403-10 (1990)). BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to the nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to the proteins of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.* (*Nucleic Acids Res.* 25(17):3389-3402 (1997)). When utilizing BLAST and gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used.

Full-length pre-processed forms, as well as mature processed forms, of proteins that comprise one of the peptides of the present invention can readily be identified as having complete sequence identity to one of the transporter peptides of the present invention as well as being encoded by the same genetic locus as the transporter peptide provided herein. As indicated by the data presented in Figure 3, the map position was determined to be on chromosome 14 by ePCR.

Allelic variants of a transporter peptide can readily be identified as being a human protein having a high degree (significant) of sequence homology/identity to at least a portion of the transporter peptide as well as being encoded by the same genetic locus as the transporter peptide provided herein. Genetic locus can readily be determined based on the genomic information provided in Figure 3, such as the genomic sequence mapped to the reference human. As indicated by the data presented in Figure 3, the map position was determined to be on chromosome 14 by ePCR. As used herein, two proteins (or a region of the proteins) have significant homology when the amino acid sequences are typically at least about 70-80%, 80-90%, and more typically at least about 90-95% or more homologous. A significantly homologous amino acid sequence, according to the present invention, will be encoded by a nucleic acid sequence that will hybridize

to a transporter peptide encoding nucleic acid molecule under stringent conditions as more fully described below.

Figure 3 provides information on SNPs that have been identified in a gene encoding the transporter protein of the present invention. 140 SNP variants were found, including 6 indels (indicated by a "-") and 1 SNPs in exons. The others were found in in introns and regions 5' and 3' of the ORF. Such SNPs in introns and outside the ORF may affect control/regulatory elements.

Paralogs of a transporter peptide can readily be identified as having some degree of significant sequence homology/identity to at least a portion of the transporter peptide, as being encoded by a gene from humans, and as having similar activity or function. Two proteins will typically be considered paralogs when the amino acid sequences are typically at least about 60% or greater, and more typically at least about 70% or greater homology through a given region or domain. Such paralogs will be encoded by a nucleic acid sequence that will hybridize to a transporter peptide encoding nucleic acid molecule under moderate to stringent conditions as more fully described below.

Orthologs of a transporter peptide can readily be identified as having some degree of significant sequence homology/identity to at least a portion of the transporter peptide as well as being encoded by a gene from another organism. Preferred orthologs will be isolated from mammals, preferably primates, for the development of human therapeutic targets and agents. Such orthologs will be encoded by a nucleic acid sequence that will hybridize to a transporter peptide encoding nucleic acid molecule under moderate to stringent conditions, as more fully described below, depending on the degree of relatedness of the two organisms yielding the proteins.

Non-naturally occurring variants of the transporter peptides of the present invention can readily be generated using recombinant techniques. Such variants include, but are not limited to deletions, additions and substitutions in the amino acid sequence of the transporter peptide. For example, one class of substitutions are conserved amino acid substitution. Such substitutions are those that substitute a given amino acid in a transporter peptide by another amino acid of like characteristics. Typically seen as conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu, and Ile; interchange of the hydroxyl residues Ser and Thr; exchange of the acidic residues Asp and Glu; substitution between the amide residues Asn and Gln; exchange of the basic residues Lys and Arg; and replacements among the aromatic residues Phe and Tyr. Guidance concerning which amino acid changes are likely to be phenotypically silent are found in Bowie *et al.*, *Science 247*:1306-1310 (1990).

Variant transporter peptides can be fully functional or can lack function in one or more activities, e.g. ability to bind ligand, ability to transport ligand, ability to mediate signaling, etc. Fully functional variants typically contain only conservative variation or variation in non-critical residues or in non-critical regions. Figure 2 provides the result of protein analysis and can be used to identify critical domains/regions. Functional variants can also contain substitution of similar amino acids that result in no change or an insignificant change in function. Alternatively, such substitutions may positively or negatively affect function to some degree.

Non-functional variants typically contain one or more non-conservative amino acid substitutions, deletions, insertions, inversions, or truncation or a substitution, insertion, inversion, or deletion in a critical residue or critical region.

Amino acids that are essential for function can be identified by methods known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham *et al.*, *Science 244*:1081-1085 (1989)), particularly using the results provided in Figure 2. The latter procedure introduces single alanine mutations at every residue in the molecule. The resulting mutant molecules are then tested for biological activity such as transporter activity or in assays such as an *in vitro* proliferative activity. Sites that are critical for binding partner/substrate binding can also be determined by structural analysis such as crystallization, nuclear magnetic resonance or photoaffinity labeling (Smith *et al., J. Mol. Biol. 224*:899-904 (1992); de Vos *et al. Science 255*:306-312 (1992)).

The present invention further provides fragments of the transporter peptides, in addition to proteins and peptides that comprise and consist of such fragments, particularly those comprising the residues identified in Figure 2. The fragments to which the invention pertains, however, are not to be construed as encompassing fragments that may be disclosed publicly prior to the present invention.

As used herein, a fragment comprises at least 8, 10, 12, 14, 16, or more contiguous amino acid residues from a transporter peptide. Such fragments can be chosen based on the ability to retain one or more of the biological activities of the transporter peptide or could be chosen for the ability to perform a function, e.g. bind a substrate or act as an immunogen. Particularly important fragments are biologically active fragments, peptides that are, for example, about 8 or more amino acids in length. Such fragments will typically comprise a domain or motif of the transporter peptide, e.g., active site, a transmembrane domain or a substrate-binding domain. Further, possible fragments include, but are not limited to, domain or motif containing fragments, soluble peptide fragments, and fragments containing immunogenic structures. Predicted domains and functional

sites are readily identifiable by computer programs well known and readily available to those of skill in the art (e.g., PROSITE analysis). The results of one such analysis are provided in Figure 2.

Polypeptides often contain amino acids other than the 20 amino acids commonly referred to as the 20 naturally occurring amino acids. Further, many amino acids, including the terminal amino acids, may be modified by natural processes, such as processing and other post-translational modifications, or by chemical modification techniques well known in the art. Common modifications that occur naturally in transporter peptides are described in basic texts, detailed monographs, and the research literature, and they are well known to those of skill in the art (some of these features are identified in Figure 2).

Known modifications include, but are not limited to, acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent crosslinks, formation of cystine, formation of pyroglutamate, formylation, gamma carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination.

Such modifications are well known to those of skill in the art and have been described in great detail in the scientific literature. Several particularly common modifications, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation, for instance, are described in most basic texts, such as *Proteins - Structure and Molecular Properties*, 2nd Ed., T.E. Creighton, W. H. Freeman and Company, New York (1993). Many detailed reviews are available on this subject, such as by Wold, F., *Posttranslational Covalent Modification of Proteins*, B.C. Johnson, Ed., Academic Press, New York 1-12 (1983); Seifter *et al.* (*Meth. Enzymol. 182*: 626-646 (1990)) and Rattan *et al.* (*Ann. N.Y. Acad. Sci. 663*:48-62 (1992)).

Accordingly, the transporter peptides of the present invention also encompass derivatives or analogs in which a substituted amino acid residue is not one encoded by the genetic code, in which a substituent group is included, in which the mature transporter peptide is fused with another compound, such as a compound to increase the half-life of the transporter peptide (for example, polyethylene glycol), or in which the additional amino acids are fused to the mature transporter peptide, such as a leader or secretory sequence or a sequence for purification of the mature transporter peptide or a pro-protein sequence.

Protein/Peptide Uses

The proteins of the present invention can be used in substantial and specific assays related to the functional information provided in the Figures; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its binding partner or ligand) in biological fluids; and as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state). Where the protein binds or potentially binds to another protein or ligand (such as, for example, in a transporter-effector protein interaction or transporter-ligand interaction), the protein can be used to identify the binding partner/ligand so as to develop a system to identify inhibitors of the binding interaction. Any or all of these uses are capable of being developed into reagent grade or kit format for commercialization as commercial products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

The potential uses of the peptides of the present invention are based primarily on the source of the protein as well as the class/action of the protein. For example, transporters isolated from humans and their human/mammalian orthologs serve as targets for identifying agents for use in mammalian therapeutic applications, e.g. a human drug, particularly in modulating a biological or pathological response in a cell or tissue that expresses the transporter. Experimental data as provided in Figure 1 indicates that sodium/calcium exchanger proteins of the present invention are expressed in humans in the heart, retina, kidney, fetal brain, and fetal heart. Specifically, a virtual northern blot shows expression in the fetal brain. In addition, PCR-based tissue screening panel indicates expression in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. A large percentage of pharmaceutical agents are being developed that modulate the activity of transporter proteins, particularly members of the sodium/calcium exchanger subfamily (see Background of the Invention). The structural and functional information provided in the Background and Figures provide specific and substantial uses for the molecules of the present invention, particularly in combination with the expression information provided in Figure 1. Experimental data as provided in Figure 1 indicates expression in humans in

brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. Such uses can readily be determined using the information provided herein, that known in the art and routine experimentation.

The proteins of the present invention (including variants and fragments that may have been disclosed prior to the present invention) are useful for biological assays related to transporters that are related to members of the sodium/calcium exchanger subfamily. Such assays involve any of the known transporter functions or activities or properties useful for diagnosis and treatment of transporter-related conditions that are specific for the subfamily of transporters that the one of the present invention belongs to, particularly in cells and tissues that express the transporter. Experimental data as provided in Figure 1 indicates that sodium/calcium exchanger proteins of the present invention are expressed in humans in the heart, retina, kidney, fetal brain, and fetal heart. Specifically, a virtual northern blot shows expression in the fetal brain. In addition, PCR-based tissue screening panel indicates expression in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain.

The proteins of the present invention are also useful in drug screening assays, in cell-based or cell-free systems ((Hodgson, Bio/technology, 1992, Sept 10(9);973-80). Cell-based systems can be native, i.e., cells that normally express the transporter, as a biopsy or expanded in cell culture. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. In an alternate embodiment, cell-based assays involve recombinant host cells expressing the transporter protein.

The polypeptides can be used to identify compounds that modulate transporter activity of the protein in its natural state or an altered form that causes a specific disease or pathology associated with the transporter. Both the transporters of the present invention and appropriate variants and fragments can be used in high-throughput screens to assay candidate compounds for the ability to bind to the transporter. These compounds can be further screened against a functional transporter to determine the effect of the compound on the transporter activity. Further, these compounds can be tested in animal or invertebrate systems to determine activity/effectiveness. Compounds can be identified that activate (agonist) or inactivate (antagonist) the transporter to a desired degree.

Further, the proteins of the present invention can be used to screen a compound for the ability to stimulate or inhibit interaction between the transporter protein and a molecule that normally interacts with the transporter protein, e.g. a substrate or a component of the signal pathway that the transporter protein normally interacts (for example, another transporter). Such assays

typically include the steps of combining the transporter protein with a candidate compound under conditions that allow the transporter protein, or fragment, to interact with the target molecule, and to detect the formation of a complex between the protein and the target or to detect the biochemical consequence of the interaction with the transporter protein and the target, such as any of the associated effects of signal transduction such as changes in membrane potential, protein phosphorylation, cAMP turnover, and adenylate cyclase activation, etc.

Candidate compounds include, for example, 1) peptides such as soluble peptides, including Ig-tailed fusion peptides and members of random peptide libraries (see, e.g., Lam *et al.*, *Nature* *354*:82-84 (1991); Houghten *et al.*, *Nature 354*:84-86 (1991)) and combinatorial chemistry-derived molecular libraries made of D- and/or L- configuration amino acids; 2) phosphopeptides (e.g., members of random and partially degenerate, directed phosphopeptide libraries, see, e.g., Songyang *et al.*, *Cell 72*:767-778 (1993)); 3) antibodies (e.g., polyclonal, monoclonal, humanized, anti-idiotypic, chimeric, and single chain antibodies as well as Fab, $F(ab')_2$, Fab expression library fragments, and epitope-binding fragments of antibodies); and 4) small organic and inorganic molecules (e.g., molecules obtained from combinatorial and natural product libraries).

One candidate compound is a soluble fragment of the receptor that competes for ligand binding. Other candidate compounds include mutant transporters or appropriate fragments containing mutations that affect transporter function and thus compete for ligand. Accordingly, a fragment that competes for ligand, for example with a higher affinity, or a fragment that binds ligand but does not allow release, is encompassed by the invention.

The invention further includes other end point assays to identify compounds that modulate (stimulate or inhibit) transporter activity. The assays typically involve an assay of events in the signal transduction pathway that indicate transporter activity. Thus, the transport of a ligand, change in cell membrane potential, activation of a protein, a change in the expression of genes that are up- or down-regulated in response to the transporter protein dependent signal cascade can be assayed.

Any of the biological or biochemical functions mediated by the transporter can be used as an endpoint assay. These include all of the biochemical or biochemical/biological events described herein, in the references cited herein, incorporated by reference for these endpoint assay targets, and other functions known to those of ordinary skill in the art or that can be readily identified using the information provided in the Figures, particularly Figure 2. Specifically, a biological function of a cell or tissues that expresses the transporter can be assayed. Experimental data as provided in Figure 1 indicates that sodium/calcium exchanger proteins of the present invention are expressed in

humans in the heart, retina, kidney, fetal brain, and fetal heart. Specifically, a virtual northern blot shows expression in the fetal brain. In addition, PCR-based tissue screening panel indicates expression in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain.

Binding and/or activating compounds can also be screened by using chimeric transporter proteins in which the amino terminal extracellular domain, or parts thereof, the entire transmembrane domain or subregions, such as any of the seven transmembrane segments or any of the intracellular or extracellular loops and the carboxy terminal intracellular domain, or parts thereof, can be replaced by heterologous domains or subregions. For example, a ligand-binding region can be used that interacts with a different ligand then that which is recognized by the native transporter. Accordingly, a different set of signal transduction components is available as an end-point assay for activation. This allows for assays to be performed in other than the specific host cell from which the transporter is derived.

The proteins of the present invention are also useful in competition binding assays in methods designed to discover compounds that interact with the transporter (e.g. binding partners and/or ligands). Thus, a compound is exposed to a transporter polypeptide under conditions that allow the compound to bind or to otherwise interact with the polypeptide. Soluble transporter polypeptide is also added to the mixture. If the test compound interacts with the soluble transporter polypeptide, it decreases the amount of complex formed or activity from the transporter target. This type of assay is particularly useful in cases in which compounds are sought that interact with specific regions of the transporter. Thus, the soluble polypeptide that competes with the target transporter region is designed to contain peptide sequences corresponding to the region of interest.

To perform cell free drug screening assays, it is sometimes desirable to immobilize either the transporter protein, or fragment, or its target molecule to facilitate separation of complexes from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay.

Techniques for immobilizing proteins on matrices can be used in the drug screening assays. In one embodiment, a fusion protein can be provided which adds a domain that allows the protein to be bound to a matrix. For example, glutathione-S-transferase fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the cell lysates (e.g., ^{35}S-labeled) and the candidate compound, and the mixture incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads are washed to remove any unbound label, and the matrix immobilized and radiolabel determined directly, or in the

supernatant after the complexes are dissociated. Alternatively, the complexes can be dissociated from the matrix, separated by SDS-PAGE, and the level of transporter-binding protein found in the bead fraction quantitated from the gel using standard electrophoretic techniques. For example, either the polypeptide or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin using techniques well known in the art. Alternatively, antibodies reactive with the protein but which do not interfere with binding of the protein to its target molecule can be derivatized to the wells of the plate, and the protein trapped in the wells by antibody conjugation. Preparations of a transporter-binding protein and a candidate compound are incubated in the transporter protein-presenting wells and the amount of complex trapped in the well can be quantitated. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the transporter protein target molecule, or which are reactive with transporter protein and compete with the target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the target molecule.

Agents that modulate one of the transporters of the present invention can be identified using one or more of the above assays, alone or in combination. It is generally preferable to use a cell-based or cell free system first and then confirm activity in an animal or other model system. Such model systems are well known in the art and can readily be employed in this context.

Modulators of transporter protein activity identified according to these drug screening assays can be used to treat a subject with a disorder mediated by the transporter pathway, by treating cells or tissues that express the transporter. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. These methods of treatment include the steps of administering a modulator of transporter activity in a pharmaceutical composition to a subject in need of such treatment, the modulator being identified as described herein.

In yet another aspect of the invention, the transporter proteins can be used as "bait proteins" in a two-hybrid assay or three-hybrid assay (see, e.g., U.S. Patent No. 5,283,317; Zervos *et al.* (1993) *Cell* 72:223-232; Madura *et al.* (1993) *J. Biol. Chem.* 268:12046-12054; Bartel *et al.* (1993) *Biotechniques* 14:920-924; Iwabuchi *et al.* (1993) *Oncogene* 8:1693-1696; and Brent WO94/10300), to identify other proteins, which bind to or interact with the transporter and are involved in transporter activity. Such transporter-binding proteins are also likely to be involved in the propagation of signals by the transporter proteins or transporter targets as, for

example, downstream elements of a transporter-mediated signaling pathway. Alternatively, such transporter-binding proteins are likely to be transporter inhibitors.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for a transporter protein is fused to a gene encoding the DNA binding domain of a known transcription factor (e.g., GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo,* forming a transporter-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (e.g., LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with the transporter protein.

This invention further pertains to novel agents identified by the above-described screening assays. Accordingly, it is within the scope of this invention to further use an agent identified as described herein in an appropriate animal model. For example, an agent identified as described herein (e.g., a transporter-modulating agent, an antisense transporter nucleic acid molecule, a transporter-specific antibody, or a transporter-binding partner) can be used in an animal or other model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal or other model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

The transporter proteins of the present invention are also useful to provide a target for diagnosing a disease or predisposition to disease mediated by the peptide. Accordingly, the invention provides methods for detecting the presence, or levels of, the protein (or encoding mRNA) in a cell, tissue, or organism. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. The method involves contacting a biological sample with a compound capable of interacting with the transporter protein such that the interaction can be detected. Such an assay can be provided in a single detection format or a multi-detection format such as an antibody chip array.

prophylactic or therapeutic treatment based on the individual's genotype. The discovery of genetic polymorphisms in some drug metabolizing enzymes has explained why some patients do not obtain the expected drug effects, show an exaggerated drug effect, or experience serious toxicity from standard drug dosages. Polymorphisms can be expressed in the phenotype of the extensive metabolizer and the phenotype of the poor metabolizer. Accordingly, genetic polymorphism may lead to allelic protein variants of the transporter protein in which one or more of the transporter functions in one population is different from those in another population. The peptides thus allow a target to ascertain a genetic predisposition that can affect treatment modality. Thus, in a ligand-based treatment, polymorphism may give rise to amino terminal extracellular domains and/or other ligand-binding regions that are more or less active in ligand binding, and transporter activation. Accordingly, ligand dosage would necessarily be modified to maximize the therapeutic effect within a given population containing a polymorphism. As an alternative to genotyping, specific polymorphic peptides could be identified.

The peptides are also useful for treating a disorder characterized by an absence of, inappropriate, or unwanted expression of the protein. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. Accordingly, methods for treatment include the use of the transporter protein or fragments.

Antibodies

The invention also provides antibodies that selectively bind to one of the peptides of the present invention, a protein comprising such a peptide, as well as variants and fragments thereof. As used herein, an antibody selectively binds a target peptide when it binds the target peptide and does not significantly bind to unrelated proteins. An antibody is still considered to selectively bind a peptide even if it also binds to other proteins that are not substantially homologous with the target peptide so long as such proteins share homology with a fragment or domain of the peptide target of the antibody. In this case, it would be understood that antibody binding to the peptide is still selective despite some degree of cross-reactivity.

As used herein, an antibody is defined in terms consistent with that recognized within the art: they are multi-subunit proteins produced by a mammalian organism in response to an antigen challenge. The antibodies of the present invention include polyclonal antibodies and monoclonal antibodies, as well as fragments of such antibodies, including, but not limited to, Fab or $F(ab')_2$, and Fv fragments.

Many methods are known for generating and/or identifying antibodies to a given target peptide. Several such methods are described by Harlow, Antibodies, Cold Spring Harbor Press, (1989).

In general, to generate antibodies, an isolated peptide is used as an immunogen and is administered to a mammalian organism, such as a rat, rabbit or mouse. The full-length protein, an antigenic peptide fragment or a fusion protein can be used. Particularly important fragments are those covering functional domains, such as the domains identified in Figure 2, and domain of sequence homology or divergence amongst the family, such as those that can readily be identified using protein alignment methods and as presented in the Figures.

Antibodies are preferably prepared from regions or discrete fragments of the transporter proteins. Antibodies can be prepared from any region of the peptide as described herein. However, preferred regions will include those involved in function/activity and/or transporter/binding partner interaction. Figure 2 can be used to identify particularly important regions while sequence alignment can be used to identify conserved and unique sequence fragments.

An antigenic fragment will typically comprise at least 8 contiguous amino acid residues. The antigenic peptide can comprise, however, at least 10, 12, 14, 16 or more amino acid residues. Such fragments can be selected on a physical property, such as fragments correspond to regions that are located on the surface of the protein, e.g., hydrophilic regions or can be selected based on sequence uniqueness (see Figure 2).

Detection on an antibody of the present invention can be facilitated by coupling (i.e., physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S or ^{3}H.

Antibody Uses

The antibodies can be used to isolate one of the proteins of the present invention by standard techniques, such as affinity chromatography or immunoprecipitation. The antibodies can facilitate the purification of the natural protein from cells and recombinantly produced protein expressed in host cells. In addition, such antibodies are useful to detect the presence of one of the proteins of the present invention in cells or tissues to determine the pattern of expression of the protein among various tissues in an organism and over the course of normal development. Experimental data as provided in Figure 1 indicates that sodium/calcium exchanger proteins of the present invention are expressed in humans in the heart, retina, kidney, fetal brain, and fetal heart. Specifically, a virtual northern blot shows expression in the fetal brain. In addition, PCR-based tissue screening panel indicates expression in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. Further, such antibodies can be used to detect protein *in situ*, *in vitro*, or in a cell lysate or supernatant in order to evaluate the abundance and pattern of expression. Also, such antibodies can be used to assess abnormal tissue distribution or abnormal expression during development or progression of a biological condition. Antibody detection of circulating fragments of the full length protein can be used to identify turnover.

Further, the antibodies can be used to assess expression in disease states such as in active stages of the disease or in an individual with a predisposition toward disease related to the protein's function. When a disorder is caused by an inappropriate tissue distribution, developmental expression, level of expression of the protein, or expressed/processed form, the antibody can be prepared against the normal protein. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. If a disorder is characterized by a specific mutation in the protein, antibodies specific for this mutant protein can be used to assay for the presence of the specific mutant protein.

The antibodies can also be used to assess normal and aberrant subcellular localization of cells in the various tissues in an organism. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. The diagnostic uses can be applied, not only in genetic testing, but also in monitoring a treatment modality. Accordingly, where treatment is ultimately aimed at correcting expression level or the presence of aberrant sequence and aberrant tissue distribution or developmental expression, antibodies directed against the protein or relevant fragments can be used to monitor therapeutic efficacy.

Additionally, antibodies are useful in pharmacogenomic analysis. Thus, antibodies prepared against polymorphic proteins can be used to identify individuals that require modified treatment modalities. The antibodies are also useful as diagnostic tools as an immunological marker for aberrant protein analyzed by electrophoretic mobility, isoelectric point, tryptic peptide digest, and other physical assays known to those in the art.

The antibodies are also useful for tissue typing. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. Thus, where a specific protein has been correlated with expression in a specific tissue, antibodies that are specific for this protein can be used to identify a tissue type.

The antibodies are also useful for inhibiting protein function, for example, blocking the binding of the transporter peptide to a binding partner such as a ligand or protein binding partner. These uses can also be applied in a therapeutic context in which treatment involves inhibiting the protein's function. An antibody can be used, for example, to block binding, thus modulating (agonizing or antagonizing) the peptides activity. Antibodies can be prepared against specific fragments containing sites required for function or against intact protein that is associated with a cell or cell membrane. See Figure 2 for structural information relating to the proteins of the present invention.

The invention also encompasses kits for using antibodies to detect the presence of a protein in a biological sample. The kit can comprise antibodies such as a labeled or labelable antibody and a compound or agent for detecting protein in a biological sample; means for determining the amount of protein in the sample; means for comparing the amount of protein in the sample with a standard; and instructions for use. Such a kit can be supplied to detect a single protein or epitope or can be configured to detect one of a multitude of epitopes, such as in an antibody detection array. Arrays are described in detail below for nucleic acid arrays and similar methods have been developed for antibody arrays.

Nucleic Acid Molecules

The present invention further provides isolated nucleic acid molecules that encode a transporter peptide or protein of the present invention (cDNA, transcript and genomic sequence). Such nucleic acid molecules will consist of, consist essentially of, or comprise a nucleotide sequence that encodes one of the transporter peptides of the present invention, an allelic variant thereof, or an ortholog or paralog thereof.

As used herein, an "isolated" nucleic acid molecule is one that is separated from other nucleic acid present in the natural source of the nucleic acid. Preferably, an "isolated" nucleic acid is free of sequences that naturally flank the nucleic acid (i.e., sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. However, there can be some flanking nucleotide sequences, for example up to about 5KB, 4KB, 3KB, 2KB, or 1KB or less, particularly contiguous peptide encoding sequences and peptide encoding sequences within the same gene but separated by introns in the genomic sequence. The important point is that the nucleic acid is isolated from remote and unimportant flanking sequences such that it can be subjected to the specific manipulations described herein such as recombinant expression, preparation of probes and primers, and other uses specific to the nucleic acid sequences.

Moreover, an "isolated" nucleic acid molecule, such as a transcript/cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or chemical precursors or other chemicals when chemically synthesized. However, the nucleic acid molecule can be fused to other coding or regulatory sequences and still be considered isolated.

For example, recombinant DNA molecules contained in a vector are considered isolated. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells or purified (partially or substantially) DNA molecules in solution. Isolated RNA molecules include *in vivo* or *in vitro* RNA transcripts of the isolated DNA molecules of the present invention. Isolated nucleic acid molecules according to the present invention further include such molecules produced synthetically.

Accordingly, the present invention provides nucleic acid molecules that consist of the nucleotide sequence shown in Figure 1 or 3 (SEQ ID NO:1, transcript sequence and SEQ ID NO:3, genomic sequence), or any nucleic acid molecule that encodes the protein provided in Figure 2, SEQ ID NO:2. A nucleic acid molecule consists of a nucleotide sequence when the nucleotide sequence is the complete nucleotide sequence of the nucleic acid molecule.

The present invention further provides nucleic acid molecules that consist essentially of the nucleotide sequence shown in Figure 1 or 3 (SEQ ID NO:1, transcript sequence and SEQ ID NO:3, genomic sequence), or any nucleic acid molecule that encodes the protein provided in Figure 2, SEQ ID NO:2. A nucleic acid molecule consists essentially of a nucleotide sequence when such a nucleotide sequence is present with only a few additional nucleic acid residues in the final nucleic acid molecule.

The present invention further provides nucleic acid molecules that comprise the nucleotide sequences shown in Figure 1 or 3 (SEQ ID NO:1, transcript sequence and SEQ ID NO:3, genomic sequence), or any nucleic acid molecule that encodes the protein provided in Figure 2, SEQ ID NO:2. A nucleic acid molecule comprises a nucleotide sequence when the nucleotide sequence is at least part of the final nucleotide sequence of the nucleic acid molecule. In such a fashion, the nucleic acid molecule can be only the nucleotide sequence or have additional nucleic acid residues, such as nucleic acid residues that are naturally associated with it or heterologous nucleotide sequences. Such a nucleic acid molecule can have a few additional nucleotides or can comprise several hundred or more additional nucleotides. A brief description of how various types of these nucleic acid molecules can be readily made/isolated is provided below.

In Figures 1 and 3, both coding and non-coding sequences are provided. Because of the source of the present invention, humans genomic sequence (Figure 3) and cDNA/transcript sequences (Figure 1), the nucleic acid molecules in the Figures will contain genomic intronic sequences, 5' and 3' non-coding sequences, gene regulatory regions and non-coding intergenic sequences. In general such sequence features are either noted in Figures 1 and 3 or can readily be identified using computational tools known in the art. As discussed below, some of the non-coding regions, particularly gene regulatory elements such as promoters, are useful for a variety of purposes, e.g. control of heterologous gene expression, target for identifying gene activity modulating compounds, and are particularly claimed as fragments of the genomic sequence provided herein.

The isolated nucleic acid molecules can encode the mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids interior to the mature peptide (when the mature form has more than one peptide chain, for instance). Such sequences may play a role in processing of a protein from precursor to a mature form, facilitate protein trafficking, prolong or shorten protein half-life or facilitate manipulation of a protein for assay or production, among other things. As generally is the case *in situ*, the additional amino acids may be processed away from the mature protein by cellular enzymes.

As mentioned above, the isolated nucleic acid molecules include, but are not limited to, the sequence encoding the transporter peptide alone, the sequence encoding the mature peptide and additional coding sequences, such as a leader or secretory sequence (e.g., a pre-pro or pro-protein sequence), the sequence encoding the mature peptide, with or without the additional coding sequences, plus additional non-coding sequences, for example introns and non-coding 5' and 3' sequences such as transcribed but non-translated sequences that play a role in transcription, mRNA

processing (including splicing and polyadenylation signals), ribosome binding and stability of mRNA. In addition, the nucleic acid molecule may be fused to a marker sequence encoding, for example, a peptide that facilitates purification.

Isolated nucleic acid molecules can be in the form of RNA, such as mRNA, or in the form DNA, including cDNA and genomic DNA obtained by cloning or produced by chemical synthetic techniques or by a combination thereof. The nucleic acid, especially DNA, can be double-stranded or single-stranded. Single-stranded nucleic acid can be the coding strand (sense strand) or the non-coding strand (anti-sense strand).

The invention further provides nucleic acid molecules that encode fragments of the peptides of the present invention as well as nucleic acid molecules that encode obvious variants of the transporter proteins of the present invention that are described above. Such nucleic acid molecules may be naturally occurring, such as allelic variants (same locus), paralogs (different locus), and orthologs (different organism), or may be constructed by recombinant DNA methods or by chemical synthesis. Such non-naturally occurring variants may be made by mutagenesis techniques, including those applied to nucleic acid molecules, cells, or organisms. Accordingly, as discussed above, the variants can contain nucleotide substitutions, deletions, inversions and insertions. Variation can occur in either or both the coding and non-coding regions. The variations can produce both conservative and non-conservative amino acid substitutions.

The present invention further provides non-coding fragments of the nucleic acid molecules provided in Figures 1 and 3. Preferred non-coding fragments include, but are not limited to, promoter sequences, enhancer sequences, gene modulating sequences and gene termination sequences. Such fragments are useful in controlling heterologous gene expression and in developing screens to identify gene-modulating agents. A promoter can readily be identified as being 5' to the ATG start site in the genomic sequence provided in Figure 3.

A fragment comprises a contiguous nucleotide sequence greater than 12 or more nucleotides. Further, a fragment could at least 30, 40, 50, 100, 250 or 500 nucleotides in length. The length of the fragment will be based on its intended use. For example, the fragment can encode epitope bearing regions of the peptide, or can be useful as DNA probes and primers. Such fragments can be isolated using the known nucleotide sequence to synthesize an oligonucleotide probe. A labeled probe can then be used to screen a cDNA library, genomic DNA library, or mRNA to isolate nucleic acid corresponding to the coding region. Further, primers can be used in PCR reactions to clone specific regions of gene.

A probe/primer typically comprises substantially a purified oligonucleotide or oligonucleotide pair. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, 20, 25, 40, 50 or more consecutive nucleotides.

Orthologs, homologs, and allelic variants can be identified using methods well known in the art. As described in the Peptide Section, these variants comprise a nucleotide sequence encoding a peptide that is typically 60-70%, 70-80%, 80-90%, and more typically at least about 90-95% or more homologous to the nucleotide sequence shown in the Figure sheets or a fragment of this sequence. Such nucleic acid molecules can readily be identified as being able to hybridize under moderate to stringent conditions, to the nucleotide sequence shown in the Figure sheets or a fragment of the sequence. Allelic variants can readily be determined by genetic locus of the encoding gene. As indicated by the data presented in Figure 3, the map position was determined to be on chromosome 14 by ePCR.

Figure 3 provides information on SNPs that have been identified in a gene encoding the transporter protein of the present invention. 140 SNP variants were found, including 6 indels (indicated by a "-") and 1 SNPs in exons. The others were found in in introns and regions 5' and 3' of the ORF. Such SNPs in introns and outside the ORF may affect control/regulatory elements.

As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences encoding a peptide at least 60-70% homologous to each other typically remain hybridized to each other. The conditions can be such that sequences at least about 60%, at least about 70%, or at least about 80% or more homologous to each other typically remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. One example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-65C. Examples of moderate to low stringency hybridization conditions are well known in the art.

<u>Nucleic Acid Molecule Uses</u>

The nucleic acid molecules of the present invention are useful for probes, primers, chemical intermediates, and in biological assays. The nucleic acid molecules are useful as a hybridization probe for messenger RNA, transcript/cDNA and genomic DNA to isolate full-length cDNA and

genomic clones encoding the peptide described in Figure 2 and to isolate cDNA and genomic clones that correspond to variants (alleles, orthologs, etc.) producing the same or related peptides shown in Figure 2. 140 SNPs, including 6 indels, have been identified in the gene encoding the transporter protein provided by the present invention and are given in Figure 3.

The probe can correspond to any sequence along the entire length of the nucleic acid molecules provided in the Figures. Accordingly, it could be derived from 5' noncoding regions, the coding region, and 3' noncoding regions. However, as discussed, fragments are not to be construed as encompassing fragments disclosed prior to the present invention.

The nucleic acid molecules are also useful as primers for PCR to amplify any given region of a nucleic acid molecule and are useful to synthesize antisense molecules of desired length and sequence.

The nucleic acid molecules are also useful for constructing recombinant vectors. Such vectors include expression vectors that express a portion of, or all of, the peptide sequences. Vectors also include insertion vectors, used to integrate into another nucleic acid molecule sequence, such as into the cellular genome, to alter *in situ* expression of a gene and/or gene product. For example, an endogenous coding sequence can be replaced via homologous recombination with all or part of the coding region containing one or more specifically introduced mutations.

The nucleic acid molecules are also useful for expressing antigenic portions of the proteins.

The nucleic acid molecules are also useful as probes for determining the chromosomal positions of the nucleic acid molecules by means of *in situ* hybridization methods. As indicated by the data presented in Figure 3, the map position was determined to be on chromosome 14 by ePCR.

The nucleic acid molecules are also useful in making vectors containing the gene regulatory regions of the nucleic acid molecules of the present invention.

The nucleic acid molecules are also useful for designing ribozymes corresponding to all, or a part, of the mRNA produced from the nucleic acid molecules described herein.

The nucleic acid molecules are also useful for making vectors that express part, or all, of the peptides.

The nucleic acid molecules are also useful for constructing host cells expressing a part, or all, of the nucleic acid molecules and peptides.

The nucleic acid molecules are also useful for constructing transgenic animals expressing all, or a part, of the nucleic acid molecules and peptides.

The nucleic acid molecules are also useful as hybridization probes for determining the presence, level, form and distribution of nucleic acid expression. Experimental data as provided in

Figure 1 indicates that sodium/calcium exchanger proteins of the present invention are expressed in humans in the heart, retina, kidney, fetal brain, and fetal heart. Specifically, a virtual northern blot shows expression in the fetal brain. In addition, PCR-based tissue screening panel indicates expression in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain.

Accordingly, the probes can be used to detect the presence of, or to determine levels of, a specific nucleic acid molecule in cells, tissues, and in organisms. The nucleic acid whose level is determined can be DNA or RNA. Accordingly, probes corresponding to the peptides described herein can be used to assess expression and/or gene copy number in a given cell, tissue, or organism. These uses are relevant for diagnosis of disorders involving an increase or decrease in transporter protein expression relative to normal results.

In vitro techniques for detection of mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detecting DNA include Southern hybridizations and *in situ* hybridization.

Probes can be used as a part of a diagnostic test kit for identifying cells or tissues that express a transporter protein, such as by measuring a level of a transporter-encoding nucleic acid in a sample of cells from a subject e.g., mRNA or genomic DNA, or determining if a transporter gene has been mutated. Experimental data as provided in Figure 1 indicates that sodium/calcium exchanger proteins of the present invention are expressed in humans in the heart, retina, kidney, fetal brain, and fetal heart. Specifically, a virtual northern blot shows expression in the fetal brain. In addition, PCR-based tissue screening panel indicates expression in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain.

Nucleic acid expression assays are useful for drug screening to identify compounds that modulate transporter nucleic acid expression.

The invention thus provides a method for identifying a compound that can be used to treat a disorder associated with nucleic acid expression of the transporter gene, particularly biological and pathological processes that are mediated by the transporter in cells and tissues that express it. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. The method typically includes assaying the ability of the compound to modulate the expression of the transporter nucleic acid and thus identifying a compound that can be used to treat a disorder characterized by undesired transporter nucleic acid expression. The assays can be performed in cell-based and cell-free systems. Cell-based assays include cells naturally expressing the transporter nucleic acid or recombinant cells genetically engineered to express specific nucleic acid sequences.

The assay for transporter nucleic acid expression can involve direct assay of nucleic acid levels, such as mRNA levels, or on collateral compounds involved in the signal pathway. Further, the expression of genes that are up- or down-regulated in response to the transporter protein signal pathway can also be assayed. In this embodiment the regulatory regions of these genes can be operably linked to a reporter gene such as luciferase.

Thus, modulators of transporter gene expression can be identified in a method wherein a cell is contacted with a candidate compound and the expression of mRNA determined. The level of expression of transporter mRNA in the presence of the candidate compound is compared to the level of expression of transporter mRNA in the absence of the candidate compound. The candidate compound can then be identified as a modulator of nucleic acid expression based on this comparison and be used, for example to treat a disorder characterized by aberrant nucleic acid expression. When expression of mRNA is statistically significantly greater in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of nucleic acid expression. When nucleic acid expression is statistically significantly less in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of nucleic acid expression.

The invention further provides methods of treatment, with the nucleic acid as a target, using a compound identified through drug screening as a gene modulator to modulate transporter nucleic acid expression in cells and tissues that express the transporter. Experimental data as provided in Figure 1 indicates that sodium/calcium exchanger proteins of the present invention are expressed in humans in the heart, retina, kidney, fetal brain, and fetal heart. Specifically, a virtual northern blot shows expression in the fetal brain. In addition, PCR-based tissue screening panel indicates expression in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain. Modulation includes both up-regulation (i.e. activation or agonization) or down-regulation (suppression or antagonization) or nucleic acid expression.

Alternatively, a modulator for transporter nucleic acid expression can be a small molecule or drug identified using the screening assays described herein as long as the drug or small molecule inhibits the transporter nucleic acid expression in the cells and tissues that express the protein. Experimental data as provided in Figure 1 indicates expression in humans in brain, heart, kidney, lung, spleen, testis, leukocyte and fetal brain.

The nucleic acid molecules are also useful for monitoring the effectiveness of modulating compounds on the expression or activity of the transporter gene in clinical trials or in a treatment regimen. Thus, the gene expression pattern can serve as a barometer for the continuing

effectiveness of treatment with the compound, particularly with compounds to which a patient can develop resistance. The gene expression pattern can also serve as a marker indicative of a physiological response of the affected cells to the compound. Accordingly, such monitoring would allow either increased administration of the compound or the administration of alternative compounds to which the patient has not become resistant. Similarly, if the level of nucleic acid expression falls below a desirable level, administration of the compound could be commensurately decreased.

The nucleic acid molecules are also useful in diagnostic assays for qualitative changes in transporter nucleic acid expression, and particularly in qualitative changes that lead to pathology. The nucleic acid molecules can be used to detect mutations in transporter genes and gene expression products such as mRNA. The nucleic acid molecules can be used as hybridization probes to detect naturally occurring genetic mutations in the transporter gene and thereby to determine whether a subject with the mutation is at risk for a disorder caused by the mutation. Mutations include deletion, addition, or substitution of one or more nucleotides in the gene, chromosomal rearrangement, such as inversion or transposition, modification of genomic DNA, such as aberrant methylation patterns or changes in gene copy number, such as amplification. Detection of a mutated form of the transporter gene associated with a dysfunction provides a diagnostic tool for an active disease or susceptibility to disease when the disease results from overexpression, underexpression, or altered expression of a transporter protein.

Individuals carrying mutations in the transporter gene can be detected at the nucleic acid level by a variety of techniques. Figure 3 provides information on SNPs that have been identified in a gene encoding the transporter protein of the present invention. 140 SNP variants were found, including 6 indels (indicated by a "-") and 1 SNPs in exons. The others were found in in introns and regions 5' and 3' of the ORF. Such SNPs in introns and outside the ORF may affect control/regulatory elements. As indicated by the data presented in Figure 3, the map position was determined to be on chromosome 14 by ePCR. Genomic DNA can be analyzed directly or can be amplified by using PCR prior to analysis. RNA or cDNA can be used in the same way. In some uses, detection of the mutation involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, e.g. U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, e.g., Landegran *et al.*, *Science 241*:1077-1080 (1988); and Nakazawa *et al.*, *PNAS 91*:360-364 (1994)), the latter of which can be particularly useful for detecting point mutations in the gene (see Abravaya *et al.*, *Nucleic Acids Res. 23*:675-682 (1995)). This method can include the steps of collecting a sample of cells from a patient, isolating

nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a gene under conditions such that hybridization and amplification of the gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. Deletions and insertions can be detected by a change in size of the amplified product compared to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to normal RNA or antisense DNA sequences.

Alternatively, mutations in a transporter gene can be directly identified, for example, by alterations in restriction enzyme digestion patterns determined by gel electrophoresis.

Further, sequence-specific ribozymes (U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site. Perfectly matched sequences can be distinguished from mismatched sequences by nuclease cleavage digestion assays or by differences in melting temperature.

Sequence changes at specific locations can also be assessed by nuclease protection assays such as RNase and S1 protection or the chemical cleavage method. Furthermore, sequence differences between a mutant transporter gene and a wild-type gene can be determined by direct DNA sequencing. A variety of automated sequencing procedures can be utilized when performing the diagnostic assays (Naeve, C.W., (1995) *Biotechniques 19*:448), including sequencing by mass spectrometry (see, e.g., PCT International Publication No. WO 94/16101; Cohen *et al.*, *Adv. Chromatogr. 36*:127-162 (1996); and Griffin *et al.*, *Appl. Biochem. Biotechnol. 38*:147-159 (1993)).

Other methods for detecting mutations in the gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA duplexes (Myers *et al.*, *Science 230*:1242 (1985)); Cotton *et al.*, *PNAS 85*:4397 (1988); Saleeba *et al.*, *Meth. Enzymol. 217*:286-295 (1992)), electrophoretic mobility of mutant and wild type nucleic acid is compared (Orita *et al.*, *PNAS 86*:2766 (1989); Cotton *et al.*, *Mutat. Res. 285*:125-144 (1993); and Hayashi *et al.*, *Genet. Anal. Tech. Appl. 9*:73-79 (1992)), and movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (Myers *et al.*, *Nature 313*:495 (1985)). Examples of other techniques for detecting point mutations include selective oligonucleotide hybridization, selective amplification, and selective primer extension.

The nucleic acid molecules are also useful for testing an individual for a genotype that while not necessarily causing the disease, nevertheless affects the treatment modality. Thus, the nucleic acid molecules can be used to study the relationship between an individual's genotype and the

individual's response to a compound used for treatment (pharmacogenomic relationship). Accordingly, the nucleic acid molecules described herein can be used to assess the mutation content of the transporter gene in an individual in order to select an appropriate compound or dosage regimen for treatment. Figure 3 provides information on SNPs that have been identified in a gene encoding the transporter protein of the present invention. 140 SNP variants were found, including 6 indels (indicated by a "-") and 1 SNPs in exons. The others were found in in introns and regions 5' and 3' of the ORF. Such SNPs in introns and outside the ORF may affect control/regulatory elements.

Thus nucleic acid molecules displaying genetic variations that affect treatment provide a diagnostic target that can be used to tailor treatment in an individual. Accordingly, the production of recombinant cells and animals containing these polymorphisms allow effective clinical design of treatment compounds and dosage regimens.

The nucleic acid molecules are thus useful as antisense constructs to control transporter gene expression in cells, tissues, and organisms. A DNA antisense nucleic acid molecule is designed to be complementary to a region of the gene involved in transcription, preventing transcription and hence production of transporter protein. An antisense RNA or DNA nucleic acid molecule would hybridize to the mRNA and thus block translation of mRNA into transporter protein.

Alternatively, a class of antisense molecules can be used to inactivate mRNA in order to decrease expression of transporter nucleic acid. Accordingly, these molecules can treat a disorder characterized by abnormal or undesired transporter nucleic acid expression. This technique involves cleavage by means of ribozymes containing nucleotide sequences complementary to one or more regions in the mRNA that attenuate the ability of the mRNA to be translated. Possible regions include coding regions and particularly coding regions corresponding to the catalytic and other functional activities of the transporter protein, such as ligand binding.

The nucleic acid molecules also provide vectors for gene therapy in patients containing cells that are aberrant in transporter gene expression. Thus, recombinant cells, which include the patient's cells that have been engineered *ex vivo* and returned to the patient, are introduced into an individual where the cells produce the desired transporter protein to treat the individual.

The invention also encompasses kits for detecting the presence of a transporter nucleic acid in a biological sample. Experimental data as provided in Figure 1 indicates that sodium/calcium exchanger proteins of the present invention are expressed in humans in the heart, retina, kidney, fetal brain, and fetal heart. Specifically, a virtual northern blot shows expression in the fetal brain. In addition, PCR-based tissue screening panel indicates expression in brain, heart, kidney, lung,

spleen, testis, leukocyte and fetal brain. For example, the kit can comprise reagents such as a labeled or labelable nucleic acid or agent capable of detecting transporter nucleic acid in a biological sample; means for determining the amount of transporter nucleic acid in the sample; and means for comparing the amount of transporter nucleic acid in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect transporter protein mRNA or DNA.

Nucleic Acid Arrays

The present invention further provides nucleic acid detection kits, such as arrays or microarrays of nucleic acid molecules that are based on the sequence information provided in Figures 1 and 3 (SEQ ID NOS:1 and 3).

As used herein "Arrays" or "Microarrays" refers to an array of distinct polynucleotides or oligonucleotides synthesized on a substrate, such as paper, nylon or other type of membrane, filter, chip, glass slide, or any other suitable solid support. In one embodiment, the microarray is prepared and used according to the methods described in US Patent 5,837,832, Chee *et al.*, PCT application W095/11995 (Chee *et al.*), Lockhart, D. J. *et al.* (1996; Nat. Biotech. 14: 1675-1680) and Schena, M. *et al.* (1996; Proc. Natl. Acad. Sci. 93: 10614-10619), all of which are incorporated herein in their entirety by reference. In other embodiments, such arrays are produced by the methods described by Brown *et al.*, US Patent No. 5,807,522.

The microarray or detection kit is preferably composed of a large number of unique, single-stranded nucleic acid sequences, usually either synthetic antisense oligonucleotides or fragments of cDNAs, fixed to a solid support. The oligonucleotides are preferably about 6-60 nucleotides in length, more preferably 15-30 nucleotides in length, and most preferably about 20-25 nucleotides in length. For a certain type of microarray or detection kit, it may be preferable to use oligonucleotides that are only 7-20 nucleotides in length. The microarray or detection kit may contain oligonucleotides that cover the known 5', or 3', sequence, sequential oligonucleotides that cover the full length sequence; or unique oligonucleotides selected from particular areas along the length of the sequence. Polynucleotides used in the microarray or detection kit may be oligonucleotides that are specific to a gene or genes of interest.

In order to produce oligonucleotides to a known sequence for a microarray or detection kit, the gene(s) of interest (or an ORF identified from the contigs of the present invention) is typically examined using a computer algorithm which starts at the 5' or at the 3' end of the nucleotide sequence. Typical algorithms will then identify oligomers of defined length that are

unique to the gene, have a GC content within a range suitable for hybridization, and lack predicted secondary structure that may interfere with hybridization. In certain situations it may be appropriate to use pairs of oligonucleotides on a microarray or detection kit. The "pairs" will be identical, except for one nucleotide that preferably is located in the center of the sequence. The second oligonucleotide in the pair (mismatched by one) serves as a control. The number of oligonucleotide pairs may range from two to one million. The oligomers are synthesized at designated areas on a substrate using a light-directed chemical process. The substrate may be paper, nylon or other type of membrane, filter, chip, glass slide or any other suitable solid support.

In another aspect, an oligonucleotide may be synthesized on the surface of the substrate by using a chemical coupling procedure and an ink jet application apparatus, as described in PCT application W095/251116 (Baldeschweiler *et al.*) which is incorporated herein in its entirety by reference. In another aspect, a "gridded" array analogous to a dot (or slot) blot may be used to arrange and link cDNA fragments or oligonucleotides to the surface of a substrate using a vacuum system, thermal, UV, mechanical or chemical bonding procedures. An array, such as those described above, may be produced by hand or by using available devices (slot blot or dot blot apparatus), materials (any suitable solid support), and machines (including robotic instruments), and may contain 8, 24, 96, 384, 1536, 6144 or more oligonucleotides, or any other number between two and one million which lends itself to the efficient use of commercially available instrumentation.

In order to conduct sample analysis using a microarray or detection kit, the RNA or DNA from a biological sample is made into hybridization probes. The mRNA is isolated, and cDNA is produced and used as a template to make antisense RNA (aRNA). The aRNA is amplified in the presence of fluorescent nucleotides, and labeled probes are incubated with the microarray or detection kit so that the probe sequences hybridize to complementary oligonucleotides of the microarray or detection kit. Incubation conditions are adjusted so that hybridization occurs with precise complementary matches or with various degrees of less complementarity. After removal of nonhybridized probes, a scanner is used to determine the levels and patterns of fluorescence. The scanned images are examined to determine degree of complementarity and the relative abundance of each oligonucleotide sequence on the microarray or detection kit. The biological samples may be obtained from any bodily fluids (such as blood, urine, saliva, phlegm, gastric juices, etc.), cultured cells, biopsies, or other tissue preparations. A detection system may be used to measure the absence, presence, and amount of hybridization for all of the distinct

sequences simultaneously. This data may be used for large-scale correlation studies on the sequences, expression patterns, mutations, variants, or polymorphisms among samples.

Using such arrays, the present invention provides methods to identify the expression of the transporter proteins/peptides of the present invention. In detail, such methods comprise incubating a test sample with one or more nucleic acid molecules and assaying for binding of the nucleic acid molecule with components within the test sample. Such assays will typically involve arrays comprising many genes, at least one of which is a gene of the present invention and or alleles of the transporter gene of the present invention. Figure 3 provides information on SNPs that have been identified in a gene encoding the transporter protein of the present invention. 140 SNP variants were found, including 6 indels (indicated by a "-") and 1 SNPs in exons. The others were found in in introns and regions 5' and 3' of the ORF. Such SNPs in introns and outside the ORF may affect control/regulatory elements.

Conditions for incubating a nucleic acid molecule with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid molecule used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization, amplification or array assay formats can readily be adapted to employ the novel fragments of the Human genome disclosed herein. Examples of such assays can be found in Chard, T, *An Introduction to Radioimmunoassay and Related Techniques,* Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G. R. *et al., Techniques in Immunocytochemistry,* Academic Press, Orlando, FL Vol. 1 (1 982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., *Practice and Theory of Enzyme Immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology,* Elsevier Science Publishers, Amsterdam, The Netherlands (1985).

The test samples of the present invention include cells, protein or membrane extracts of cells. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing nucleic acid extracts or of cells are well known in the art and can be readily be adapted in order to obtain a sample that is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention.

Specifically, the invention provides a compartmentalized kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the nucleic acid molecules that can bind to a fragment of the Human genome disclosed herein; and

(b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound nucleic acid.

In detail, a compartmentalized kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers, strips of plastic, glass or paper, or arraying material such as silica. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the nucleic acid probe, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound probe. One skilled in the art will readily recognize that the previously unidentified transporter gene of the present invention can be routinely identified using the sequence information disclosed herein can be readily incorporated into one of the established kit formats which are well known in the art, particularly expression arrays.

Vectors/host cells

The invention also provides vectors containing the nucleic acid molecules described herein. The term "vector" refers to a vehicle, preferably a nucleic acid molecule, which can transport the nucleic acid molecules. When the vector is a nucleic acid molecule, the nucleic acid molecules are covalently linked to the vector nucleic acid. With this aspect of the invention, the vector includes a plasmid, single or double stranded phage, a single or double stranded RNA or DNA viral vector, or artificial chromosome, such as a BAC, PAC, YAC, OR MAC.

A vector can be maintained in the host cell as an extrachromosomal element where it replicates and produces additional copies of the nucleic acid molecules. Alternatively, the vector may integrate into the host cell genome and produce additional copies of the nucleic acid molecules when the host cell replicates.

The invention provides vectors for the maintenance (cloning vectors) or vectors for expression (expression vectors) of the nucleic acid molecules. The vectors can function in procaryotic or eukaryotic cells or in both (shuttle vectors).

Expression vectors contain cis-acting regulatory regions that are operably linked in the vector to the nucleic acid molecules such that transcription of the nucleic acid molecules is allowed in a host cell. The nucleic acid molecules can be introduced into the host cell with a separate

nucleic acid molecule capable of affecting transcription. Thus, the second nucleic acid molecule may provide a trans-acting factor interacting with the cis-regulatory control region to allow transcription of the nucleic acid molecules from the vector. Alternatively, a trans-acting factor may be supplied by the host cell. Finally, a trans-acting factor can be produced from the vector itself. It is understood, however, that in some embodiments, transcription and/or translation of the nucleic acid molecules can occur in a cell-free system.

The regulatory sequence to which the nucleic acid molecules described herein can be operably linked include promoters for directing mRNA transcription. These include, but are not limited to, the left promoter from bacteriophage λ, the lac, TRP, and TAC promoters from *E. coli*, the early and late promoters from SV40, the CMV immediate early promoter, the adenovirus early and late promoters, and retrovirus long-terminal repeats.

In addition to control regions that promote transcription, expression vectors may also include regions that modulate transcription, such as repressor binding sites and enhancers. Examples include the SV40 enhancer, the cytomegalovirus immediate early enhancer, polyoma enhancer, adenovirus enhancers, and retrovirus LTR enhancers.

In addition to containing sites for transcription initiation and control, expression vectors can also contain sequences necessary for transcription termination and, in the transcribed region a ribosome binding site for translation. Other regulatory control elements for expression include initiation and termination codons as well as polyadenylation signals. The person of ordinary skill in the art would be aware of the numerous regulatory sequences that are useful in expression vectors. Such regulatory sequences are described, for example, in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual. 2nd. ed.*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, (1989).

A variety of expression vectors can be used to express a nucleic acid molecule. Such vectors include chromosomal, episomal, and virus-derived vectors, for example vectors derived from bacterial plasmids, from bacteriophage, from yeast episomes, from yeast chromosomal elements, including yeast artificial chromosomes, from viruses such as baculoviruses, papovaviruses such as SV40, Vaccinia viruses, adenoviruses, poxviruses, pseudorabies viruses, and retroviruses. Vectors may also be derived from combinations of these sources such as those derived from plasmid and bacteriophage genetic elements, e.g. cosmids and phagemids. Appropriate cloning and expression vectors for prokaryotic and eukaryotic hosts are described in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual. 2nd. ed.*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, (1989).

The regulatory sequence may provide constitutive expression in one or more host cells (i.e. tissue specific) or may provide for inducible expression in one or more cell types such as by temperature, nutrient additive, or exogenous factor such as a hormone or other ligand. A variety of vectors providing for constitutive and inducible expression in prokaryotic and eukaryotic hosts are well known to those of ordinary skill in the art.

The nucleic acid molecules can be inserted into the vector nucleic acid by well-known methodology. Generally, the DNA sequence that will ultimately be expressed is joined to an expression vector by cleaving the DNA sequence and the expression vector with one or more restriction enzymes and then ligating the fragments together. Procedures for restriction enzyme digestion and ligation are well known to those of ordinary skill in the art.

The vector containing the appropriate nucleic acid molecule can be introduced into an appropriate host cell for propagation or expression using well-known techniques. Bacterial cells include, but are not limited to, *E. coli*, *Streptomyces*, and *Salmonella typhimurium*. Eukaryotic cells include, but are not limited to, yeast, insect cells such as *Drosophila*, animal cells such as COS and CHO cells, and plant cells.

As described herein, it may be desirable to express the peptide as a fusion protein. Accordingly, the invention provides fusion vectors that allow for the production of the peptides. Fusion vectors can increase the expression of a recombinant protein, increase the solubility of the recombinant protein, and aid in the purification of the protein by acting for example as a ligand for affinity purification. A proteolytic cleavage site may be introduced at the junction of the fusion moiety so that the desired peptide can ultimately be separated from the fusion moiety. Proteolytic enzymes include, but are not limited to, factor Xa, thrombin, and enterotransporter. Typical fusion expression vectors include pGEX (Smith *et al.*, *Gene 67*:31-40 (1988)), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein. Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann *et al.*, *Gene 69*:301-315 (1988)) and pET 11d (Studier *et al.*, *Gene Expression Technology: Methods in Enzymology 185*:60-89 (1990)).

Recombinant protein expression can be maximized in host bacteria by providing a genetic background wherein the host cell has an impaired capacity to proteolytically cleave the recombinant protein. (Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Alternatively, the sequence of the nucleic acid

molecule of interest can be altered to provide preferential codon usage for a specific host cell, for example *E. coli*. (Wada *et al.*, *Nucleic Acids Res. 20*:2111-2118 (1992)).

The nucleic acid molecules can also be expressed by expression vectors that are operative in yeast. Examples of vectors for expression in yeast e.g., *S. cerevisiae* include pYepSec1 (Baldari, *et al.*, *EMBO J. 6*:229-234 (1987)), pMFa (Kurjan *et al.*, *Cell 30*:933-943(1982)), pJRY88 (Schultz *et al.*, *Gene 54*:113-123 (1987)), and pYES2 (Invitrogen Corporation, San Diego, CA).

The nucleic acid molecules can also be expressed in insect cells using, for example, baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (e.g., Sf 9 cells) include the pAc series (Smith *et al.*, *Mol. Cell Biol. 3*:2156-2165 (1983)) and the pVL series (Lucklow *et al.*, *Virology 170*:31-39 (1989)).

In certain embodiments of the invention, the nucleic acid molecules described herein are expressed in mammalian cells using mammalian expression vectors. Examples of mammalian expression vectors include pCDM8 (Seed, B. *Nature 329*:840(1987)) and pMT2PC (Kaufman *et al.*, *EMBO J. 6*:187-195 (1987)).

The expression vectors listed herein are provided by way of example only of the well-known vectors available to those of ordinary skill in the art that would be useful to express the nucleic acid molecules. The person of ordinary skill in the art would be aware of other vectors suitable for maintenance propagation or expression of the nucleic acid molecules described herein. These are found for example in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

The invention also encompasses vectors in which the nucleic acid sequences described herein are cloned into the vector in reverse orientation, but operably linked to a regulatory sequence that permits transcription of antisense RNA. Thus, an antisense transcript can be produced to all, or to a portion, of the nucleic acid molecule sequences described herein, including both coding and non-coding regions. Expression of this antisense RNA is subject to each of the parameters described above in relation to expression of the sense RNA (regulatory sequences, constitutive or inducible expression, tissue-specific expression).

The invention also relates to recombinant host cells containing the vectors described herein. Host cells therefore include prokaryotic cells, lower eukaryotic cells such as yeast, other eukaryotic cells such as insect cells, and higher eukaryotic cells such as mammalian cells.

The recombinant host cells are prepared by introducing the vector constructs described herein into the cells by techniques readily available to the person of ordinary skill in the art. These

include, but are not limited to, calcium phosphate transfection, DEAE-dextran-mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, lipofection, and other techniques such as those found in Sambrook, *et al.* (*Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Host cells can contain more than one vector. Thus, different nucleotide sequences can be introduced on different vectors of the same cell. Similarly, the nucleic acid molecules can be introduced either alone or with other nucleic acid molecules that are not related to the nucleic acid molecules such as those providing trans-acting factors for expression vectors. When more than one vector is introduced into a cell, the vectors can be introduced independently, co-introduced or joined to the nucleic acid molecule vector.

In the case of bacteriophage and viral vectors, these can be introduced into cells as packaged or encapsulated virus by standard procedures for infection and transduction. Viral vectors can be replication-competent or replication-defective. In the case in which viral replication is defective, replication will occur in host cells providing functions that complement the defects.

Vectors generally include selectable markers that enable the selection of the subpopulation of cells that contain the recombinant vector constructs. The marker can be contained in the same vector that contains the nucleic acid molecules described herein or may be on a separate vector. Markers include tetracycline or ampicillin-resistance genes for prokaryotic host cells and dihydrofolate reductase or neomycin resistance for eukaryotic host cells. However, any marker that provides selection for a phenotypic trait will be effective.

While the mature proteins can be produced in bacteria, yeast, mammalian cells, and other cells under the control of the appropriate regulatory sequences, cell- free transcription and translation systems can also be used to produce these proteins using RNA derived from the DNA constructs described herein.

Where secretion of the peptide is desired, which is difficult to achieve with multi-transmembrane domain containing proteins such as transporters, appropriate secretion signals are incorporated into the vector. The signal sequence can be endogenous to the peptides or heterologous to these peptides.

Where the peptide is not secreted into the medium, which is typically the case with transporters, the protein can be isolated from the host cell by standard disruption procedures, including freeze thaw, sonication, mechanical disruption, use of lysing agents and the like. The peptide can then be recovered and purified by well-known purification methods including

ammonium sulfate precipitation, acid extraction, anion or cationic exchange chromatography, phosphocellulose chromatography, hydrophobic-interaction chromatography, affinity chromatography, hydroxylapatite chromatography, lectin chromatography, or high performance liquid chromatography.

It is also understood that depending upon the host cell in recombinant production of the peptides described herein, the peptides can have various glycosylation patterns, depending upon the cell, or maybe non-glycosylated as when produced in bacteria. In addition, the peptides may include an initial modified methionine in some cases as a result of a host-mediated process.

Uses of vectors and host cells

The recombinant host cells expressing the peptides described herein have a variety of uses. First, the cells are useful for producing a transporter protein or peptide that can be further purified to produce desired amounts of transporter protein or fragments. Thus, host cells containing expression vectors are useful for peptide production.

Host cells are also useful for conducting cell-based assays involving the transporter protein or transporter protein fragments, such as those described above as well as other formats known in the art. Thus, a recombinant host cell expressing a native transporter protein is useful for assaying compounds that stimulate or inhibit transporter protein function.

Host cells are also useful for identifying transporter protein mutants in which these functions are affected. If the mutants naturally occur and give rise to a pathology, host cells containing the mutations are useful to assay compounds that have a desired effect on the mutant transporter protein (for example, stimulating or inhibiting function) which may not be indicated by their effect on the native transporter protein.

Genetically engineered host cells can be further used to produce non-human transgenic animals. A transgenic animal is preferably a mammal, for example a rodent, such as a rat or mouse, in which one or more of the cells of the animal include a transgene. A transgene is exogenous DNA that is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal in one or more cell types or tissues of the transgenic animal. These animals are useful for studying the function of a transporter protein and identifying and evaluating modulators of transporter protein activity. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, and amphibians.

A transgenic animal can be produced by introducing nucleic acid into the male pronuclei of a fertilized oocyte, e.g., by microinjection, retroviral infection, and allowing the oocyte to develop

in a pseudopregnant female foster animal. Any of the transporter protein nucleotide sequences can be introduced as a transgene into the genome of a non-human animal, such as a mouse.

Any of the regulatory or other sequences useful in expression vectors can form part of the transgenic sequence. This includes intronic sequences and polyadenylation signals, if not already included. A tissue-specific regulatory sequence(s) can be operably linked to the transgene to direct expression of the transporter protein to particular cells.

Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, both by Leder *et al.*, U.S. Patent No. 4,873,191 by Wagner *et al.* and in Hogan, B., *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the transgene in its genome and/or expression of transgenic mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene can further be bred to other transgenic animals carrying other transgenes. A transgenic animal also includes animals in which the entire animal or tissues in the animal have been produced using the homologously recombinant host cells described herein.

In another embodiment, transgenic non-human animals can be produced which contain selected systems that allow for regulated expression of the transgene. One example of such a system is the *cre/loxP* recombinase system of bacteriophage P1. For a description of the *cre/loxP* recombinase system, see, e.g., Lakso *et al. PNAS 89*:6232-6236 (1992). Another example of a recombinase system is the FLP recombinase system of *S. cerevisiae* (O'Gorman *et al. Science 251*:1351-1355 (1991). If a *cre/loxP* recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the *Cre* recombinase and a selected protein is required. Such animals can be provided through the construction of "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut, I. *et al. Nature 385*:810-813 (1997) and PCT International Publication Nos. WO 97/07668 and WO 97/07669. In brief, a cell, e.g., a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G_0 phase. The quiescent cell can then be fused, e.g., through the use of electrical pulses, to an enucleated

oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyst and then transferred to pseudopregnant female foster animal. The offspring born of this female foster animal will be a clone of the animal from which the cell, e.g., the somatic cell, is isolated.

Transgenic animals containing recombinant cells that express the peptides described herein are useful to conduct the assays described herein in an *in vivo* context. Accordingly, the various physiological factors that are present *in vivo* and that could effect ligand binding, transporter protein activation, and signal transduction, may not be evident from *in vitro* cell-free or cell-based assays. Accordingly, it is useful to provide non-human transgenic animals to assay *in vivo* transporter protein function, including ligand interaction, the effect of specific mutant transporter proteins on transporter protein function and ligand interaction, and the effect of chimeric transporter proteins. It is also possible to assess the effect of null mutations, that is mutations that substantially or completely eliminate one or more transporter protein functions.

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the above-described modes for carrying out the invention which are obvious to those skilled in the field of molecular biology or related fields are intended to be within the scope of the following claims.

Claims

That which is claimed is:

1. An isolated peptide consisting of an amino acid sequence selected from the group consisting of:

(a) an amino acid sequence shown in SEQ ID NO:2;

(b) an amino acid sequence of an allelic variant of an amino acid sequence shown in SEQ ID NO:2, wherein said allelic variant is encoded by a nucleic acid molecule that hybridizes under stringent conditions to the opposite strand of a nucleic acid molecule shown in SEQ ID NOS:1 or 3;

(c) an amino acid sequence of an ortholog of an amino acid sequence shown in SEQ ID NO:2, wherein said ortholog is encoded by a nucleic acid molecule that hybridizes under stringent conditions to the opposite strand of a nucleic acid molecule shown in SEQ ID NOS:1 or 3; and

(d) a fragment of an amino acid sequence shown in SEQ ID NO:2, wherein said fragment comprises at least 10 contiguous amino acids.

2. An isolated peptide comprising an amino acid sequence selected from the group consisting of:

(a) an amino acid sequence shown in SEQ ID NO:2;

(b) an amino acid sequence of an allelic variant of an amino acid sequence shown in SEQ ID NO:2, wherein said allelic variant is encoded by a nucleic acid molecule that hybridizes under stringent conditions to the opposite strand of a nucleic acid molecule shown in SEQ ID NOS:1 or 3;

(c) an amino acid sequence of an ortholog of an amino acid sequence shown in SEQ ID NO:2, wherein said ortholog is encoded by a nucleic acid molecule that hybridizes under stringent conditions to the opposite strand of a nucleic acid molecule shown in SEQ ID NOS:1 or 3; and

(d) a fragment of an amino acid sequence shown in SEQ ID NO:2, wherein said fragment comprises at least 10 contiguous amino acids.

3. An isolated antibody that selectively binds to a peptide of claim 2.

4. An isolated nucleic acid molecule consisting of a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence that encodes an amino acid sequence shown in SEQ ID NO:2;

(b) a nucleotide sequence that encodes of an allelic variant of an amino acid sequence shown in SEQ ID NO:2, wherein said nucleotide sequence hybridizes under stringent conditions to the opposite strand of a nucleic acid molecule shown in SEQ ID NOS:1 or 3;

(c) a nucleotide sequence that encodes an ortholog of an amino acid sequence shown in SEQ ID NO:2, wherein said nucleotide sequence hybridizes under stringent conditions to the opposite strand of a nucleic acid molecule shown in SEQ ID NOS:1 or 3;

(d) a nucleotide sequence that encodes a fragment of an amino acid sequence shown in SEQ ID NO:2, wherein said fragment comprises at least 10 contiguous amino acids; and

(e) a nucleotide sequence that is the complement of a nucleotide sequence of (a)-(d).

5. An isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence that encodes an amino acid sequence shown in SEQ ID NO:2;

(b) a nucleotide sequence that encodes of an allelic variant of an amino acid sequence shown in SEQ ID NO:2, wherein said nucleotide sequence hybridizes under stringent conditions to the opposite strand of a nucleic acid molecule shown in SEQ ID NOS:1 or 3;

(c) a nucleotide sequence that encodes an ortholog of an amino acid sequence shown in SEQ ID NO:2, wherein said nucleotide sequence hybridizes under stringent conditions to the opposite strand of a nucleic acid molecule shown in SEQ ID NOS:1 or 3;

(d) a nucleotide sequence that encodes a fragment of an amino acid sequence shown in SEQ ID NO:2, wherein said fragment comprises at least 10 contiguous amino acids; and

(e) a nucleotide sequence that is the complement of a nucleotide sequence of (a)-(d).

6. A gene chip comprising a nucleic acid molecule of claim 5.

7. A transgenic non-human animal comprising a nucleic acid molecule of claim 5.

8. A nucleic acid vector comprising a nucleic acid molecule of claim 5.

9. A host cell containing the vector of claim 8.

10. A method for producing any of the peptides of claim 1 comprising introducing a nucleotide sequence encoding any of the amino acid sequences in (a)-(d) into a host cell, and culturing the host cell under conditions in which the peptides are expressed from the nucleotide sequence.

11. A method for producing any of the peptides of claim 2 comprising introducing a nucleotide sequence encoding any of the amino acid sequences in (a)-(d) into a host cell, and culturing the host cell under conditions in which the peptides are expressed from the nucleotide sequence.

12. A method for detecting the presence of any of the peptides of claim 2 in a sample, said method comprising contacting said sample with a detection agent that specifically allows detection of the presence of the peptide in the sample and then detecting the presence of the peptide.

13. A method for detecting the presence of a nucleic acid molecule of claim 5 in a sample, said method comprising contacting the sample with an oligonucleotide that hybridizes to said nucleic acid molecule under stringent conditions and determining whether the oligonucleotide binds to said nucleic acid molecule in the sample.

14. A method for identifying a modulator of a peptide of claim 2, said method comprising contacting said peptide with an agent and determining if said agent has modulated the function or activity of said peptide.

15. The method of claim 14, wherein said agent is administered to a host cell comprising an expression vector that expresses said peptide.

16. A method for identifying an agent that binds to any of the peptides of claim 2, said method comprising contacting the peptide with an agent and assaying the contacted mixture to determine whether a complex is formed with the agent bound to the peptide.

17. A pharmaceutical composition comprising an agent identified by the method of claim 16 and a pharmaceutically acceptable carrier therefor.

18. A method for treating a disease or condition mediated by a human transporter protein, said method comprising administering to a patient a pharmaceutically effective amount of an agent identified by the method of claim 16.

19. A method for identifying a modulator of the expression of a peptide of claim 2, said method comprising contacting a cell expressing said peptide with an agent, and determining if said agent has modulated the expression of said peptide.

20. An isolated human transporter peptide having an amino acid sequence that shares at least 70% homology with an amino acid sequence shown in SEQ ID NO:2.

21. A peptide according to claim 20 that shares at least 90 percent homology with an amino acid sequence shown in SEQ ID NO:2.

22. An isolated nucleic acid molecule encoding a human transporter peptide, said nucleic acid molecule sharing at least 80 percent homology with a nucleic acid molecule shown in SEQ ID NOS:1 or 3.

23. A nucleic acid molecule according to claim 22 that shares at least 90 percent homology with a nucleic acid molecule shown in SEQ ID NOS:1 or 3.

```
   1 GTCTCGTGTA TGGCGTGGTT AAGGTTGCAG CCTCTCACCT CTGCCTTCCT
  51 CCATTTTGGG CTGGTTACCT TTGTGCTCTT CCTGAATGGT CTTCGAGCAG
 101 AGGCTGGTGG CTCAGGGGAC GTGCCAAGCA CAGGGCAGAA CAATGAGTCC
 151 TGTTCAGGGT CATCGGACTG CAAGGAGGGT GTCATCCTGC CAATCTGGTA
 201 CCCGGAGAAC CCTTCCCTTG GGGACAAGAT TGCCAGGGTC ATTGTCTATT
 251 TTGTGGCCCT GATATACATG TTCCTTGGGG TGTCCATCAT TGCTGACCGC
 301 TTCATGGCAT CTATTGAAGT CATCACCTCT CAAGAGAGGG AGGTGACAAT
 351 TAAGAAACCC AATGGAGAAA CCAGCACAAC CACTATTCGG GTCTGGAATG
 401 AAACTGTCTC CAACCTGACC CTTATGGCCC TGGGTTCCTC TGCTCCTGAG
 451 ATACTCCTCT CTTTAATTGA GGTGTGTGGT CATGGGTTCA TTGCTGGTGA
 501 TCTGGGACCT TCTACCATTG TAGGGAGTGC AGCCTTCAAC ATGTTCATCA
 551 TCATTGGCAT CTGTGTCTAC GTGATCCCAG ACGGAGAGAC TCGCAAGATC
 601 AAGCATCTAC GAGTCTTCTT CATCACCGCT GCTTGGAGTA TCTTTGCCTA
 651 CATCTGGCTC TATATGATTC TGGCAGTCTT CTCCCCTGGT GTGGTCCAGG
 701 TTTGGGAAGG CCTCCTCACT CTCTTCTTCT TTCCAGTGTG TGTCCTTCTG
 751 GCCTGGGTGG CAGATAAACG ACTGCTCTTC TACAAATACA TGCACAAAAA
 801 GTACCGCACA GACAAACACC GAGGAATTAT CATAGAGACA GAGGGTGACC
 851 ACCCTAAGGG CATTGAGATG GATGGGAAAA TGATGAATTC CCATTTTCTA
 901 GATGGGAACC TGGTGCCCCT GGAAGGGAAG GAAGTGGATG AGTCCCGCAG
 951 AGAGATGATC CGGATCCTCA AGGATCTGAA GCAAAAACAC CCAGAGAAGG
1001 ACTTAGATCA GCTGGTGGAG ATGGCCAATT ACTATGCTCT TTCCCACCAA
1051 CAGAAGAGCC GCGCCTTCTA CCGTATCCAA GCCACTCGTA TGATGACTGG
1101 TGCAGGCAAT ATCCTGAAGA AACATGCAGC AGAACAAGCC AAGAAGGCCT
1151 CCAGCATGAG CGAGGTGCAC ACCGATGAGC CTGAGGACTT TATTTCCAAG
1201 GTCTTCTTTG ACCCATGTTC TTACCAGTGC CTGGAGAACT GTGGGGCTGT
1251 ACTCCTGACA GTGGTGAGGA AAGGGGGAGA CATGTCAAAG ACCATGTATG
1301 TGGACTACAA AACAGAGGAT GGTTCTGCCA ATGCAGGGGC TGACTATGAG
1351 TTCACAGAGG GCACGGTGGT TCTGAAGCCA GGAGAGACCC AGAAGGAGTT
1401 CTCCGTGGGC ATAATTGATG ACGACATTTT TGAGGAGGAT GAACACTTCT
1451 TTGTAAGGTT GAGCAATGTC CGCATAGAGG AGGAGCAGCC AGAGGAGGGG
1501 ATGCCTCCAG CAATATTCAA CAGTCTTCCC TTGCCTCGGG CTGTCCTAGC
1551 CTCCCCTTGT GTGGCCACAG TTACCATCTT GGATGATGAC CATGCAGGCA
1601 TCTTCACTTT TGAATGTGAT ACTATTCATG TCAGTGAGAG TATTGGTGTT
1651 ATGGAGGTCA AGGTTCTGCG GACATCAGGT GCCCGGGGTA CAGTCATCGT
1701 CCCCTTTAGG ACAGTAGAAG GGACAGCCAA GGGTGGCGGT GAGGACTTTG
1751 AAGACACATA TGGGGAGTTG GAATTCAAGA ATGATGAAAC TGTGAAAACC
1801 ATAAGGGTTA AAATAGTAGA TGAGGAGGAA TACGAAAGGC AAGAGAATTT
1851 CTTCATTGCC CTTGGTGAAC CGAAATGGAT GGAACGTGGA ATATCAGATG
1901 TGACAGACAG GAAGCTGACT ATGGAAGAAG AGGAGGCCAA GAGGATAGCA
1951 GAGATGGGAA AGCCAGTATT GGGTGAACAC CCCAAACTGG AAGTCATCAT
2001 TGAAGAGTCC TATGAGTTCA AGACTACGGT GGACAAACTG ATCAAGAAGA
2051 CAAACCTGGC CTTGGTTGTG GGGACCCATT CCTGGAGGGA CCAGTTCATG
2101 GAGGCCATCA CCGTCAGTGC AGCAGGGGAT GAGGATGAGG ATGAATCCGG
2151 GGAGGAGAGG CTGCCCTCCT GCTTTGACTA CGTCATGCAC TTCCTGACTG
2201 TCTTCTGGAA GGTGCTGTTT GCCTGTGTGC CCCCCACAGA GTACTGCCAC
2251 GGCTGGGCCT GCTTCGCCGT CTCCATCCTC ATCATTGGCA TGCTCACCGC
2301 CATCATTGGG GACCTGGCCT CGCACTTCGG CTGCACCATT GGTCTCAAAG
2351 ATTCGGTCAC AGCTGTTGTT TTCGTGGCAT TTGGCACCTC TGTCCCAGAT
2401 ACGTTTGCCA GCAAAGCTGC TGCCCTCCAG GATGTATATG CAGACGCCTC
2451 CATTGGCAAC GTGACGGGCA GCAACGCCGT CAATGTCTTC CTGGGCATCG
2501 GCCTGGCCTG GTCCGTGGCC GCCATCTACT GGGCTCTGCA GGGACAGGAG
2551 TTCCACGTGT CGGCCGGCAC ACTGGCCTTC TCCGTCACCC TCTTCACCAT
2601 CTTTGCATTT GTCTGCATCA GCGTGCTCTT GTACCGAAGG CGGCCGCACC
2651 TGGGAGGGGA GCTTGGTGGC CCCCGTGGCT GCAAGCTCGC CACAACATGG
2701 CTCTTTGTGA GCCTGTGGCT CCTCTACATA CTCTTTGCCA CACTAGAGGC
2751 CTATTGCTAC ATCAAGGGGT TCTAAGCCAC AC
(SEQ ID NO: 1)
```

```
5'UTR:        1 - 9
Start Codon: 10
Stop Codon:  2773
3'UTR:       2776
```

FIGURE 1, page 1 of 2

HOMOLOGOUS PROTEINS:

Top 10 BLAST Hits:

Sequences producing significant alignments:	Score (bits)	E Value
CRA|18000005047237 /altid=gi|2498054 /def=sp|P70549|NAC3_RAT SO...	1828	0.0
CRA|18000005200270 /altid=gi|4140706 /def=gb|AAD04173.1| (AF107...	1342	0.0
CRA|1000682343796 /altid=gi|6453729 /def=gb|AAF08988.1|AF108389...	1338	0.0
CRA|18000004939788 /altid=gi|1083801 /def=pir||S43730 Na+/Ca2+-...	1335	0.0
CRA|18000005028314 /altid=gi|1279782 /def=gb|AAA97928.1| (U5266...	1334	0.0
CRA|18000004968774 /altid=gi|382752 /def=prf||1901175A Na/Ca ex...	1333	0.0
CRA|18000004882912 /altid=gi|627801 /def=pir||B53335 Na+/Ca2+-e...	1331	0.0
CRA|18000005218648 /altid=gi|4566522 /def=gb|AAD23386.1|AF10916...	1330	0.0
CRA|18000005218651 /altid=gi|4566528 /def=gb|AAD23389.1|AF10916...	1329	0.0
CRA|18000004907324 /altid=gi|479177 /def=pir||S32435 Na+/Ca2+-e...	1328	0.0

dbEST:

Sequences producing significant alignments:	Score (bits)	E Value
gi|11600765 /dataset=dbest /taxon=96...	500	e-138
gi|318815 /dataset=dbest /taxon=9606 /...	216	2e-53

EXPRESSION INFORMATION FOR MODULATORY USE:

gi|11600765 Pooled (Brain, Heart, Kidney, Lung, Spleen, Testis, Leukocyte)
gi|318815 Fetal brain

Tissue expression:

Pooled tissues (Brain, Heart, Kidney, Lung, Spleen, Testis, Leukocyte)

FIGURE 1, page 2 of 2

```
  1 MAWLRLQPLT SAFLHFGLVT FVLFLNGLRA EAGGSGDVPS TGQNNESCSG
 51 SSDCKEGVIL PIWYPENPSL GDKIARVIVY FVALIYMFLG VSIIADRFMA
101 SIEVITSQER EVTIKKPNGE TSTTTIRVWN ETVSNLTLMA LGSSAPEILL
151 SLIEVCGHGF IAGDLGPSTI VGSAAFNMFI IIGICVYVIP DGETRKIKHL
201 RVFFITAAWS IFAYIWLYMI LAVFSPGVVQ VWEGLLTLFF FPVCVLLAWV
251 ADKRLLFYKY MHKKYRTDKH RGIIIETEGD HPKGIEMDGK MMNSHFLDGN
301 LVPLEGKEVD ESRREMIRIL KDLKQKHPEK DLDQLVEMAN YYALSHQQKS
351 RAFYRIQATR MMTGAGNILK KHAAEQAKKA SSMSEVHTDE PEDFISKVFF
401 DPCSYQCLEN CGAVLLTVVR KGGDMSKTMY VDYKTEDGSA NAGADYEFTE
451 GTVVLKPGET QKEFSVGIID DDIFEEDEHF FVRLSNVRIE EEQPEEGMPP
501 AIFNSLPLPR AVLASPCVAT VTILDDDHAG IFTFECDTIH VSESIGVMEV
551 KVLRTSGARG TVIVPFRTVE GTAKGGGEDF EDTYGELEFK NDETVKTIRV
601 KIVDEEEYER QENFFIALGE PKWMERGISD VTDRKLTMEE EEAKRIAEMG
651 KPVLGEHPKL EVIIEESYEF KTTVDKLIKK TNLALVVGTH SWRDQFMEAI
701 TVSAAGDEDE DESGEERLPS CFDYVMHFLT VFWKVLFACV PPTEYCHGWA
751 CFAVSILIIG MLTAIIGDLA SHFGCTIGLK DSVTAVVFVA FGTSVPDTFA
801 SKAAALQDVY ADASIGNVTG SNAVNVFLGI GLAWSVAAIY WALQGQEFHV
851 SAGTLAFSVT LFTIFAFVCI SVLLYRRRPH LGGELGGPRG CKLATTWLFV
901 SLWLLYILFA TLEAYCYIKG F
 (SEQ ID NO:2)
```

FEATURES:
Functional domains and key regions:

[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site

Number of matches: 4

1	45-48	NESC
2	130-133	NETV
3	135-138	NLTL
4	817-820	NVTG

---

[2] PDOC00004 PS00004 CAMP_PHOSPHO_SITE
cAMP- and cGMP-dependent protein kinase phosphorylation site

Number of matches: 2

1	378-381	KKAS
2	634-637	RKLT

---

[3] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 11

1	113-115	TIK
2	125-127	TIR
3	597-599	TIR
4	194-196	TRK
5	267-269	TDK
6	312-314	SRR
7	460-462	TQK
8	572-574	TAK
9	594-596	TVK
10	125-127	TIR
11	597-599	TIR

FIGURE 2, page 1 of 4

```
[4] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 16
      1      69-72 SLGD
      2    106-109 TSQE
      3    144-147 SAPE
      4    151-154 SLIE
      5    277-280 TEGD
      6    312-315 SRRE
      7    382-385 SMSE
      8    460-463 TQKE
      9    522-525 TILD
     10    583-586 TYGE
     11    637-640 TMEE
     12    672-675 TTVD
     13    691-694 SWRD
     14    713-716 SGEE
     15    720-723 SCFD
     16    794-797 SVPD
------------------------------------------------------------------------
[5] PDOC00007 PS00007 TYR_PHOSPHO_SITE
Tyrosine kinase phosphorylation site

Number of matches: 2
      1    397-405 KVFFDPCSY
      2    601-608 KIVDEEEY
------------------------------------------------------------------------
[6] PDOC00008 PS00008 MYRISTYL
N-myristoylation site

Number of matches: 15
      1      50-55 GSSDCK
      2    422-427 GGDMSK
      3    438-443 GSANAG
      4    497-502 GMPPAI
      5    557-562 GARGTV
      6    571-576 GTAKGG
      7    760-765 GMLTAI
      8    774-779 GCTIGL
      9    778-783 GLKDSV
     10    816-821 GNVTGS
     11    829-834 GIGLAW
     12    831-836 GLAWSV
     13    882-887 GGELGG
     14    886-891 GGPRGC
     15    890-895 GCKLAT
```

<u>Membrane spanning structure and domains:</u>

Helix	Begin	End	Score	Certainity
1	8	28	1.905	Certain
2	76	96	2.032	Certain
3	133	153	1.009	Certain
4	169	189	1.943	Certain
5	206	226	2.118	Certain
6	231	251	2.072	Certain
7	505	525	0.666	Putative
8	723	743	1.298	Certain
9	747	767	2.258	Certain
10	781	801	1.232	Certain
11	823	843	1.793	Certain
12	854	874	2.424	Certain
13	893	913	2.138	Certain

FIGURE 2, page 2 of 4

```
BLAST Alignment to Top Hit:
>CRA|18000005047237 /altid=gi|2498054 /def=sp|P70549|NAC3_RAT
           SODIUM/CALCIUM EXCHANGER 3 PRECURSOR (NA+/CA2+-EXCHANGE
           PROTEIN 3) /org=NA+/CA2+-EXCHANGE PROTEIN 3 /dataset=nraa
           /length=927
          Length = 927

 Score = 1828 bits (4682), Expect = 0.0
 Identities = 897/927 (96%), Positives = 911/927 (97%), Gaps = 6/927 (0%)
 Frame = +1

Query: 10   MAWLRLQPLTSAFLHFGLVTFVLFLNGLRAEAGGSGDVPSTGQNNESCSGSSDCKEGVIL 189
            MAWLRLQPLTSAFLHFGLVTFVLFLNGLRAEAG    DVPS GQNNESCSGSSDCKEGVIL
Sbjct: 1    MAWLRLQPLTSAFLHFGLVTFVLFLNGLRAEAGDLRDVPSAGQNNESCSGSSDCKEGVIL 60

Query: 190  PIWYPENPSLGDKIARVIVYFVALIYMFLGVSIIADRFMASIEVITSQEREVTIKKPNGE 369
            PIWYPENPSLGDKIARVIVYFVALIYMFLGVSIIADRFMASIEVITSQEREVTIKKPNGE
Sbjct: 61   PIWYPENPSLGDKIARVIVYFVALIYMFLGVSIIADRFMASIEVITSQEREVTIKKPNGE 120

Query: 370  TSTTTIRVWNETVSNLTLMALGSSAPEILLSLIEVCGHGFIAGDLGPSTIVGSAAFNMFI 549
            TSTTTIRVWNETVSNLTLMALGSSAPEILLSLIEVCGHGFIAGDLGPSTIVGSAAFNMFI
Sbjct: 121  TSTTTIRVWNETVSNLTLMALGSSAPEILLSLIEVCGHGFIAGDLGPSTIVGSAAFNMFI 180

Query: 550  IIGICVYVIPDGETRKIKHLRVFFITAAWSIFAYIWLYMILAVFSPGVVQVWEGLLTLFF 729
            IIGICVYVIPDGETRKIKHLRVFF+TAAWS+FAYIWLYMILAVFSPGVVQVWEGLLTLFF
Sbjct: 181  IIGICVYVIPDGETRKIKHLRVFFVTAAWSVFAYIWLYMILAVFSPGVVQVWEGLLTLFF 240

Query: 730  FPVCVLLAWVADKRLLFYKYMHKKYRTDKHRGIIIETEGDHPKGIEMDGKMMNSHFLDGN 909
            FPVCVLLAWVADKRLLFYKYMHK+YRTDKHRGIIIETEG+HPKGIEMDGKMMNSHFLDGN
Sbjct: 241  FPVCVLLAWVADKRLLFYKYMHKRYRTDKHRGIIIETEGEHPKGIEMDGKMMNSHFLDGN 300

Query: 910  LVPLEGKEVDESRREMIRILKDLKQKHPEKDLDQLVEMANYYALSHQQKSRAFYRIQATR 1089
            L+PLEGKEVDESRREMIRILKDLKQKHPEKDLDQLVEMANYYALSHQQKSRAFYRIQATR
Sbjct: 301  LIPLEGKEVDESRREMIRILKDLKQKHPEKDLDQLVEMANYYALSHQQKSRAFYRIQATR 360

Query: 1090 MMTGAGNILKKHAAEQAKKASSMSEVHTDEPEDFISKVFFDPCSYQCLENCGAVLLTVVR 1269
            MMTGAGNILKKHAAEQAKK +SMSEVHTDEPEDF SKVFFDPCSYQCLENCGAVLLTVVR
Sbjct: 361  MMTGAGNILKKHAAEQAKKTASMSEVHTDEPEDFASKVFFDPCSYQCLENCGAVLLTVVR 420

Query: 1270 KGGDMSKTMYVDYKTEDGSANAGADYEFTEGTVVLKPGETQKEFSVGIIDDDIFEEDEHF 1449
            KGGD+SKTMYVDYKTEDGSANAGADYEFTEGTVVLKPGETQKEFSVGIIDDDIFEEDEHF
Sbjct: 421  KGGDISKTMYVDYKTEDGSANAGADYEFTEGTVVLKPGETQKEFSVGIIDDDIFEEDEHF 480

Query: 1450 FVRLSNVRIEEEQPEEGMPPAIFNSLPLPRAVLASPCVATVTILDDDHAGIFTFECDTIH 1629
            FVRLSNVR+EEEQ EEGM PAI NSLPLPRAVLASPCVATVTILDDDHAGIFTFECDTIH
Sbjct: 481  FVRLSNVRVEEEQLEEGMTPAILNSLPLPRAVLASPCVATVTILDDDHAGIFTFECDTIH 540

Query: 1630 VSESIGVMEVKVLRTSGARGTVIVPFRTVEGTAKGGGEDFEDTYGELEFKNDETVKTIRV 1809
            VSESIGVMEVKVLRTSGARGTVIVPFRTVEGTAKGGGEDFEDTYGELEFKNDETVKTIRV
Sbjct: 541  VSESIGVMEVKVLRTSGARGTVIVPFRTVEGTAKGGGEDFEDTYGELEFKNDETVKTIRV 600

Query: 1810 KIVDEEEYERQENFFIALGEPKWMERGIS------DVTDRKLTMEEEEAKRIAEMGKPVL 1971
            KIVDEEEYERQENFFIALGEPKWMERGIS      +VTDRKLTMEEEEAKRIAEMGKPVL
Sbjct: 601  KIVDEEEYERQENFFIALGEPKWMERGISALLLSPEVTDRKLTMEEEEAKRIAEMGKPVL 660

Query: 1972 GEHPKLEVIIEESYEFKTTVDKLIKKTNLALVVGTHSWRDQFMEAITVSAAGDEDEDESG 2151
            GEHPKLEVIIEESYEFK+TVDKLIKKTNLALVVGTHSWRDQFMEAITVSAAGDE+EDESG
Sbjct: 661  GEHPKLEVIIEESYEFKSTVDKLIKKTNLALVVGTHSWRDQFMEAITVSAAGDEEEDESG 720

Query: 2152 EERLPSCFDYVMHFLTVFWKVLFACVPPTEYCHGWACFAVSILIIGMLTAIIGDLASHFG 2331
            EERLPSCFDYVMHFLTVFWKVLFACVPPTEYCHGWACF VSILIIGMLTAIIGDLASHFG
Sbjct: 721  EERLPSCFDYVMHFLTVFWKVLFACVPPTEYCHGWACFVVSILIIGMLTAIIGDLASHFG 780

Query: 2332 CTIGLKDSVTAVVFVAFGTSVPDTFASKAAALQDVYADASIGNVTGSNAVNVFLGIGLAW 2511
            CTIGLKDSVTAVVFVAFGTSVPDTFASKAAALQDVYADASIGNVTGSNAVNVFLGIGLAW
Sbjct: 781  CTIGLKDSVTAVVFVAFGTSVPDTFASKAAALQDVYADASIGNVTGSNAVNVFLGIGLAW 840

Query: 2512 SVAAIYWALQGQEFHVSAGTLAFSVTLFTIFAFVCISVLLYRRRPHLGGELGGPRGCKLA 2691
            SVAAIYWA+QGQEFHVSAGTLAFSVTLFTIFAFVC+SVLLYRRRPHLGGELGGPRGCKLA
Sbjct: 841  SVAAIYWAMQGQEFHVSAGTLAFSVTLFTIFAFVCLSVLLYRRRPHLGGELGGPRGCKLA 900

Query: 2692 TTWLFVSLWLLYILFATLEAYCYIKGF 2772
            TTWLFVSLWLLY+LFATLEAYCYIKGF
Sbjct: 901  TTWLFVSLWLLYVLFATLEAYCYIKGF 927 (SEQ ID NO:4)

Hmmer search results (Pfam):
```

FIGURE 2, page 3 of 4

Scores for sequence family classification (score includes all domains):

Model	Description	Score	E-value	N
PF01699	Sodium/calcium exchanger protein	294.6	1.2e-84	2
PF00324	Amino acid permease	2.8	5.9	1
PF01971	Protein of unknown function	2.7	8.7	1

Parsed for domains:

Model	Domain	seq-f	seq-t		hmm-f	hmm-t		score	E-value
PF01699	1/2	118	257	..	12	152	.]	121.3	1.8e-32
PF01971	1/1	644	670	..	193	222	..	2.7	8.7
PF00324	1/1	851	877	..	472	498	.]	2.8	5.9
PF01699	2/2	757	905	..	1	152	[]	181.4	1.5e-50

FIGURE 2, page 4 of 4

```
   1 TTGGATGAGA TCTAAAGCAT TATTAAGAGT GGGGAGTGCA AAGAAGAAAC
  51 CCTCATTTCA AAGATGAATG AGAATAATGG CATGTACAAA GGTCCTGGGG
 101 TGGACAGTCA CTTGGTATAA TCCAAGAGTG AACCTGAAGG CTATTGTTGT
 151 TGAAATGTAA TAAGGGAGAG AGTGACGGGA TGAAGGGGGA TGAGTGGGAA
 201 GCAGTGAATT CCTGCAAGGC TTTGAAGGTC ATGGGAAAGA ATTTGGTCTT
 251 TATATCAAGA GCAAGAGAAG ACTACTAAAG GGCTTCAAAC AGGGGAGCGA
 301 TATGCTTAAG TCTGTTTGTT TGTTTTTTTA AAAAAAGATT ACGGTGGCTA
 351 TATGAGGAAA GTGGAATTGA GAACTAGCGA GAGTTGGAGT GGTGAGCTCC
 401 ATTAGGAGGC TACTGAAGTA GATTCATGAG GTAAGGAGTG ATGGTGGCCT
 451 GGGCTGGGAT GATGGTGGTA GAAATGGAGA AAGAGTTGAT AGGATTTAGT
 501 GATTGGATAA GGGACAGAAG AGAGATGAAG GCTTTCAGAC TAACATCTGC
 551 TTTCTAACAT GAGTAACTGG GTGGCTGAAG ATGCTATTTT CTGAGCTGGG
 601 AAACAGGAGA AAAAGGAGCA AATATGGGGG ATGAAGACTT TGAGTCTTTA
 651 AGGTGCTGTA CAAACACAAA TCAGCATTCC TTTATTACTA AGGGTATCCC
 701 ACACAGTTGT AGCAGAGGGA GAAAGATCGC CCCCCCCCCA CTTTTTTTTT
 751 TTTTTTAGCT ATTCCATGGT ATTTTCATTC TCATCCCACC CAAATGAGGC
 801 AGTGAGTGGT AAGATGAGTA TATAATAGTT TCAATTGCAT TTCATCCCAT
 851 TCTTCTGAGC TCAAGCTCAC CTTTTAGTGG TTTGAGGCCA GTAGATGAAG
 901 CTGCATATCA CCCCCAAAAT CTTGTCTCTA GTTTAACAAA ACTTATTTGA
 951 GAGACATTTG CATGTTTTAT TAATAATGAT TTTTACCACT TGTTCCTTTC
1001 CATGTTTGGG TTTGAAATTT GAGTGGCTGG CGGATGATCA TCTTCCTGTT
1051 ACTGCCTGCT TAAACTGCTC ATAAGCAGGT TTTACTGGAG GGCTCAGAGC
1101 TGCTGTGAAC TTGGTCTTGG GCACAACTTA CATGGCCTCT GTTTGGCTAT
1151 GGGGTGGGTG GCATTCACCA TTTATCAACT CTTTTGATTT CCCAAGCTAT
1201 CTCAGAATTA TAGCTTGCCT CCAGAAGTCT TGCATTCGGG GAGGAAGTTT
1251 CTTTCCAAGG GAGCTCAGTT TTCAAGGTTT ATTGCTCTGT TTAATGGATG
1301 AGATCTAAAG CATTATTAAG AGTGGGGAGT GCAAAGAAGA AACACTCATT
1351 TCAAAATCGA TTGAGAATAA TGGCATGTAC AAAGGTCCTG GGGTGGACAG
1401 TCACTTGGTA TAATCCTGGA GTGAACATGA AGGCCAAGGA AATATGTATA
1451 CATTAAACAG AGCAAGGTTT TCAATTTTCT GGGGACTAGT CCATGAAAAT
1501 TCAATTCAAT ATACTCTCTT GCAAACCTAT GTTATCCAAG ATACTCAAGT
1551 ATAATGACAA CAGGGTAAGG AAGTCCGAAC ACCCCAGAAA CAGTATAAAT
1601 GGGCATGAAG ATTCAGGTTA TACATGGCCT ATTTTAAGTT GCTTCTTGAG
1651 AACTCTCACA GGTAATACCA GTTTGGGAGA CAGGACTTGA AGGCTATTGC
1701 TGCATTTCCA TCCCCAGTAT TCCCAGCTAT TTCAAGCCAT TTTTCAACGG
1751 AGTCTCCACC AGATGGTTTG GAGGACAGAG CAGCTATTTG TGCCTCCCAT
1801 TGACATCTAT TTTTCCAAGT GAGAGACTGC CCCATATGTT AGTGCAATAT
1851 GTCACTGGAG GTGAAGCATC AGTTGTATTG GTGGGAACCT GCCGTTTGCT
1901 GTCCCCTTTT TCCTCATGCC TTTTCCTGCC TCTCTGATCT TTTCTAGGTC
1951 TCTGGCCTAT CAGGAGGACA ACTGGTGCTG CAATAGAAGC CAGTGGCTAA
2001 GTCTCGTGTA TGGCGTGGTT AAGGTTGCAG CCTCTCACCT CTGCCTTCCT
2051 CCATTTTGGG CTGGTTACCT TTGTGCTCTT CCTGAATGGT CTTCGAGCAG
2101 AGGCTGGTGG CTCAGGGGAC GTGCCAAGCA CAGGGCAGAA CAATGAGTCC
2151 TGTTCAGGGT CATCGGACTG CAAGGAGGGT GTCATCCTGC CAATCTGGTA
2201 CCCGGAGAAC CCTTCCCTTG GGGACAAGAT TGCCAGGGTC ATTGTCTATT
2251 TTGTGGCCCT GATATACATG TTCCTTGGGG TGTCCATCAT TGCTGACCGC
2301 TTCATGGCAT CTATTGAAGT CATCACCTCT CAAGAGAGGG AGGTGACAAT
2351 TAAGAAACCC AATGGAGAAA CCAGCACAAC AACTATTCGG GTCTGGAATG
2401 AAACTGTCTC CAACCTGACC CTTATGGCCC TGGGTTCCTC TGCTCCTGAG
2451 ATACTCCTCT CTTTAATTGA GGTGTGTGGT CATGGGTTCA TTGCTGGTGA
2501 TCTGGGACCT TCTACCATTG TAGGGAGTGC AGCCTTCAAC ATGTTCATCA
2551 TCATTGGCAT CTGTGTCTAC GTGATCCCAG ACGGAGAGAC TCGCAAGATC
2601 AAGCATCTAC GAGTCTTCTT CATCACCGCT GCTTGGAGTA TCTTTGCCTA
2651 CATCTGGCTC TATATGATTC TGGCAGTCTT CTCCCCTGGT GTGGTCCAGG
2701 TTTGGGAAGG CCTCCTCACT CTCTTCTTCT TTCCAGTGTG TGTCCTTCTG
2751 GCCTGGGTGG CAGATAAACG ACTGCTCTTC TACAAATACA TGCACAAAAA
2801 GTACCGCACA GACAAACACC GAGGAATTAT CATAGAGACA GAGGGTGACC
2851 ACCCTAAGGG CATTGAGATG GATGGGAAAA TGATGAATTC CCATTTTCTA
2901 GATGGGAACC TGGTGCCCCT GGAAGGGAAG GAAGTGGATG AGTCCCGCAG
2951 AGAGATGATC CGGATTCTCA AGGATCTGAA GCAAAAACAC CCAGAGAAGG
3001 ACTTAGATCA GCTGGTGGAG ATGGCCAATT ACTATGCTCT TTCCCACCAA
3051 CAGAAGAGCC GCGCCTTCTA CCGTATCCAA GCCACTCGTA TGATGACTGG
3101 TGCAGGCAAT ATCCTGAAGA AACATGCAGC AGAACAAGCC AAGAAGGCCT
3151 CCAGCATGAG CGAGGTGCAC ACCGATGAGC CTGAGGACTT TATTTCCAAG
3201 GTCTTCTTTG ACCCATGTTC TTACCAGTGC CTGGAGAACT GTGGGGCTGT
3251 ACTCCTGACA GTGGTGAGGA AAGGGGGAGA CATGTCAAAG ACCATGTATG
3301 TGGACTACAA AACAGAGGAT GGTTCTGCCA ATGCAGGGGC TGACTATGAG
3351 TTCACAGAGG GCACGGTGGT TCTGAAGCCA GGAGAGACCC AGAAGGAGTT
3401 CTCCGTGGGC ATAATTGATG ACGACATTTT TGAGGAGGAT GAACACTTCT
3451 TTGTAAGGTT GAGCAATGTC CGCATAGAGG AGGAGCAGCC AGAGGAGGGG
3501 ATGCCTCCAG CAATATTCAA CAGTCTTCCC TTGCCTCGGG CTGTCCTAGC
3551 CTCCCCTTGT GTGGCCACAG TTACCATCTT GGATGATGAC CATGCAGGCA
3601 TCTTCACTTT TGAATGTGAT ACTATTCATG TCAGTGAGAG TATTGGTGTT
3651 ATGGAGGTCA AGGTTCTGCG GACATCAGGT GCCCGGGGTA CAGTCATCGT
3701 CCCCTTTAGG ACAGTAGAAG GGACAGCCAA GGGTGGCGGT GAGGACTTTG
3751 AAGACACATA TGGGGAGTTG GAATTCAAGA ATGATGAAAC TGTGTAAGTA
3801 ACCTTCCTGT ATTCTGCCCC TCCCTGACCC CATCTTTTGC CATCTCTTTC
```

FIGURE 3, page 1 of 57

```
3851 TGTCTTTCTG TACTGCACTT TACAACATTT CCTTGTGTTT GTGTTAATGT
3901 CAAACTTTGG TTCCATCACA GGTATGCAGG ATCAGCAGAC ACCACTGGAC
3951 AGGTTCTGCT TCCAAACTCT TCTTCAGTTT TCTCACTTTA AATTGTTTCT
4001 GGGCAAGGAA TCCTGTGACA AGAGCTAAGG ACACAAAACA TTTTCTTCTC
4051 TGAAACACAA AATGATAGCT GGTGGAGCTG TGGGATGACA GAAGTTTTGT
4101 GATATCAGAT TTTGGAGAAT TCTTGTGACT AAGAAGGACT AGAGAACTGC
4151 TTGGGCCTCT TCTTCCTCCC TTCCTCATAT GAAGGGTATC TATGAGCTTT
4201 GAAACCAATC CTTTCCATTC TGGGCAGCAA TAGCCCATCA GAACATTCTA
4251 AAGAAAACAA GTGGCATTGG CTTTGTTCCC TGGTACTATA TTGCCAGTCT
4301 CACTGTGTAA CCAGATTCCA GGCACGTCTT CTTTAATTTG GAAATTGCAA
4351 AATTGATAGA AATTTAGCAA TCTTTTTAAA TGACCATAGA CTATTTAATG
4401 GTGTGAGGCT TGCCCAGCCT AGTTGAATTG AGTCAGTATG GTTTGGATAC
4451 TGGAAAGTAT CTTGGAGAAG CAGAGCTCCC AGGGCAGTGG CTACTTGTCT
4501 TTAGTCACAG GTCTAAGCTC CAAAATCTGG TGAAGCAGTG AAGGAGAAAC
4551 ATCCTAGGAA TTGTGGGAGG AAATATATCT TCTGTGTGGT CCTCTCTTTT
4601 CACAGTCTAG GACTCTCCTG AAGTACCTCT TCTTGGGCTA CTGCCCCATT
4651 CAGCCCTTCA GAAACTGTGG GTATTACACT TCTGTCACCT CTATTACCCT
4701 AAGGCCTCTG CCCATTGAAC CCTCTTGCAA ATTGGTTATT CTGTCCTTTT
4751 TCCAGTTGGA TAGCTTTAAA AGGGAAAGCA GAATGACTTT CCTCAGGATT
4801 TGTAGCTTAT GAGAAAGTAG ACTTTCTTGG GTGGCCTAGA AGGTTGGAGA
4851 AGACAAACGG GAACTTCCTC TGAATGACTG AACATATCCA CAAATAATAA
4901 GCGTGGCAGG AGATGGTGTG AAGAGTAAAA GGAGCATATA GGAAGTTGTG
4951 TGTGTGGGGT GTCTGTTTCA AGAACCTGCT AATTATACCT TCAGTAAGAA
5001 ATGAAGCCAT ACAACCTCTA GAAGAGGAGG AGGAAGGAAC TCATGGAAAA
5051 GTGGGGAGCC ATAGAAGCTA GGGAGAGGTG TCCTAGGAGT GCTTCTGCCC
5101 AGGTCCAGCC ATGAGACAGA GCTCAAAAAG AGCTGGGCAC TGCTGGTGAC
5151 AGAACTGAGT GACCCGGGGG ATCCTGCATC TGTTCTTACT CAATCCCTTC
5201 TTAATAATGT GACTTGGGGC AGGTCATTTA TTGGTTCTGG AACTTAACTT
5251 TCTGATATGC AAACTGGGAA TAACAATACT TTCCTTGCCT GGAGGCAAGG
5301 TCAGTCCTTT TTGCAGTTCC TTCCAGCTCT AAGATTTTCT GAACCATAGA
5351 CATAAGCACT CAGTGTAGGT CATATTCGCA CTTGCCAAAA ATGGATCAGG
5401 GAATATTGTC TCCTGAAGGG AAATGGCCAT TGACAAATTG ATTTATTAGA
5451 GCTCTGTTTA GTCATTTTGC TGGGAAGGAT AATCATTTGT TAACGTAAGT
5501 AGAAACCTGT GCCTTCTGGA GAATACTATC CATTTATATG TACTCTGGGG
5551 AGAGTGTTTA TACATACAAA TGAAGGACAG GGCTTCACTG GGAAAACAAA
5601 CTCCATGGAA TTTCACATGA TTATCGCGAT GTCAGTGTGG AAGAAGATAT
5651 GGTAAGGCAT TAAATGACAT TAAGACCACA AAATTTGCCA TAATTTGACG
5701 GACTTGTGGT TCTTCTGATT CAGAACCCTT TCTACCCATG TCACGGATAG
5751 GTAGTTTTTC AGAGATCAGA GGCTTAGTTC ATTCTATTAA TTTCCTCATT
5801 CTATTAATAA TCAATTATGC ACCTAGGGTC TCTGAATACG ACTAAACCTT
5851 CCTCAAACTT ATTTGCATTT TCAGTTTGTA TAATATCTTG GTGCAAATGA
5901 GCCTCGCAAA TGATCACTTC TGGGTAATAC TCATTCTAAA GGTATGTCAA
5951 CCTTGAGAAT TCTGGTCTAG ATATTCTAGG GTTTGGTGAA CAAATCTATG
6001 TTCCCATCCA TCCCTTTTCA TTTATTTTTT AGACTTCATT CATTGCAGAA
6051 TAATGAGTCC AAAACCTGCT CATCTGTTCT CACGTGGCAC CCCTATTCTT
6101 GATATTTTAA ATTGCAATTT TACAACTAGA GGCAGTATTA CGGAGCAGAA
6151 AAATCGTGGG TTCTAAGTAC TCTGGGTTAG GATTCTGGCT CCACTACTGA
6201 TTTAATAATG TAGTTTGGGG AAATTTTATT AACCTATGAA ATTATTTCCT
6251 CATTGGCAAA ATGGGGATAA TAATATCTCT CTTGCAGGGC CATTATGACG
6301 ATTCAAGGTA TTGTATGCGG TGTACCTGGT ACACGGTATA TGCTCAGGAA
6351 ACAAGACTCT TCATAGTAAT ATTGACGAAT TAACAATATT CTTCAGAAGA
6401 CACTGTGGAG TTGTTTAGGT TACTTGGCTC TTTGTGTGAC CCTAAGTAAT
6451 GAGCATGCCA GTTTGGGGTT ACTATGAAGA GTACTTACCT AAACTCATAA
6501 AATATTAGAG CTAGAAAGGA CCTTAGAATA TCTTCTGCAG TCATGGTTCT
6551 TAAATTTTAA TGTGTTGCTC AATCATCCAG GGATCTCACT GAAGGGCAGA
6601 TTAGGATCCA GGAGGTCTAG GGGAGGGATT GAGATTCCGC ATTTCTAACA
6651 AGTTCTGGAT GCTGCGGGCC CCAACTTAGA GGTGAAAGGT TCTGAAGCTC
6701 TTGACCAAAC CAGGAGACCC AGCAAAGAAG TGGTTTTTCA GACAACTTGC
6751 TTAATTGAAT AATGATTGTT TGCTCTTTAA TTCCAACTTT CAATGCCAAT
6801 TTAGCAAGAA CCAGAGGCTG TGCTAATTGC CACACCAGTC TGGAAACCGA
6851 AATGGATAGC TTCAGGGTAC TTGGACAAAG TTGGAACATC TGCTTTCTAA
6901 TCTCTCCCTC TTTGTATAGC TTTATTTGCC TACCAAGCCT GGTAGTATTG
6951 AAAATCTGCC CTCACTATAC TCCCCTAAAT ATAATCAAGT TGAGGCCAGG
7001 CCTGTGCTCT ATCAATAATA TAGGATCCAC GAATTCACAT GTTTGGTTTT
7051 ATGCTTTACT TCTTCAAAGG TGCTTTTAGC AGCATGGAAG AATGGAAAAG
7101 CACGAGCTTT GGAATATGAA AGCAGATGTG AATCCATCAC TTACCAGTAA
7151 CTTTTAACAA GTCACATCAC TTTTCTGAGT ACCAGGTTTT TGTTGGACAA
7201 CAGAAATAAT ATTCTCTATC CTTCAAGGGA ATACTAAATA TAAGTATGAG
7251 AAAAATGCAC AGTGCCTTCT CGTAGATGGT GTTCAGTCAT TCAACAAACA
7301 TTTGTTAGAT ATTTGCTATG TACTAGCTAC ATTACTAGGC ACTGGGGTTA
7351 AATAAGTGAA TAAGACAAGC TGACATTTCA GCGCTCAAGG ATCTTACTGT
7401 CAAGTGGAGA GGATCAAAGG GTACAGACAA ATCAAGGAAC GTGAGAGAAG
7451 TGGTATGGCT GAGATGGATT GAATAAAGGA GCAATGAGAG CTCCCTGCAA
7501 TGTGTGTGGT ACCACTGAGG ATTCTAAATT AACCTTCATT AAGGACTTAG
7551 TAGTGACAGA GGTGAAGTGG GGATAGGTAC ATGATTAATT TACATCCATA
7601 TTACAATGAA ACCTTAACAT TTAAGAGGGA TATTATTGAT GTCTTCATGA
7651 TCCAGAAGAA TCCTCACCTT TGCAACCATC ACTATAGTCA CTTCTTGAGA
```

FIGURE 3, page 2 of 57

7701 ATTATGGCCT TTAAGACTGT AGCATGCAAT GACAAAACCT CACAGAGGTA
7751 TGGGTTCTGC CCGCACACTA ATTTCACTCA TTAAACAAGT GACTGGCTCC
7801 TATATCCCAG GCTCTCAGCA CGCCTTTGCA AAATAACAGA TTATTGCAGC
7851 TCTTGGACCT TTGATGCCTC TGGGAATAGT CAAAGCCACA GATGTCAAAT
7901 ATGTAAATGC CAAGATCTAT TATAATTAAA TAGTGCAGGC CTCCTTCAAA
7951 GAAAAAAAGC ATGTTGGCTG TGCTGCACGT TCTCCAACCA AATCAGAATG
8001 TTAAAGCTCG AAGGTATCTG ACCTCCCATT TTTTAAATTA TGAAGATGAA
8051 ATTCAGAAAG GGAAGGTAAC TTATCCAAGA TTACATGGCT AGCTATGATA
8101 GAAAGTTAGA GTTGGAAAGG ACGTTAGAAA GTGAGGGTTT GAAAGGACTT
8151 TAGAAGCTGC TTATTCAATG TTCTCTCTGC CCTTTCCCAT CTTAGGCTTC
8201 TCCATTTTAC TTTTATCCAT CAATAAAATG TTAACTTCAA AAAGAATATG
8251 GCAATTCTTG GGTAAAAGAT GCTCTGGAAG TGTGAGTCCG GGAGTATTAT
8301 GTGACTAATG TCTTAACTAA GAATAATAAT ATATTATGGA CTAGTTTTAA
8351 TCTCTTGTTT CACCTTGAAC TGTTCAGGAA GGAAAATAGC CCACGGAAAT
8401 TTTTTAAAAA GTCTTTCTCT ATCTGAATTG AGAAAAGGTG ACAGGCATAG
8451 TTGGAACATC TTTTAGGCAG TGCTGGTGAA CTTCAGGCTA GGCCTTGTTC
8501 CATGAAATAA TAAAAATTTT CAAAATAATG CAGACCATTC CCTTCCAGGG
8551 ATGCTTTCTC TGTAATGTTT TAACCCCAAG AAATCTTTCT GTAAAAATCT
8601 ATAAAAATCT GGAGTGTTCC AGGATACAAT TTGCACATTC TCCAATTTAA
8651 CTAAAACACA ATCGATTTTT TGTTTTCTTT TTCTTTGGCT TAGCAAGGTT
8701 TTAAGATAGT CTCTTTCTGG CCACAGAGGG AGATGATTTG CCTCTAGAAT
8751 ACCCTTTCTG TGCTTGAGAG AGTCACAAGA CTGCAAGCTC ATGGAGGATG
8801 AGAGTCAAGT AGAGGTGGTG ACATCTCTCC CTTGGCCAAC ATCCCTCTCT
8851 TTCTCTTTCC TTCTGCCTTC AGTGGCAGTA GCAAAAGTCC TCCTTCTCTT
8901 TAGGTAGACA GTCAGCCACT ACAACTGTGG CTTCCTGAAA TCCTCAGTGG
8951 AGCTATGTAC TTGGCACAGA TTTGTCTTGA AGAAGGGACT CCATTTCTGA
9001 GCCAGTTGTT GAATGGGGAT ACTTAGCAGT ACAGTGAGGC ATTTCCAGTA
9051 GGATTGTTCA ACCACAATTG CCCACTTTCC AGGCCCAAAG GAATAATTGA
9101 AGGCTATGTA GACTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTGAGATGG
9151 AGTCTCGCTC TGTCGCCCAG GCTGGAGTGC AGTGGCACAT CTCGGCTCAC
9201 TGCAAGCTCT GCCTCCCGGG TTCACGCCAT TCTCCTGCCT CAGCCTCCCG
9251 AGTAGCTAGG CCTAATATAT ATATATTATA CATATATATT TATATTTATA
9301 TATATATATA CCACCACGTC CGGCTAATAT ATATTTATAC TTTTTTTTTT
9351 TAGTAGGAAA GGGGTTTCAC CATGTTAGCC AGTATGGTCT CGATCTCCTG
9401 ACCTCGTGAT CCACCAGCCT CAGCCTCCCA AAGTGCTGGG ATTACAGGCG
9451 TGAGCCACCG TGCCCGACCA TGCTATGTAA ACTTTTTAGC AGAAGCTTTA
9501 GCTATTGTGT CCCGAAGGGC CCCAGGTCAT GATGAAATGT CTTTTTTTTT
9551 TTTTGTCTCT TTTCTTCTTA ATTACTGAGA CTGTCAAAGA ATATGTCAAA
9601 GCATGACATA TTCCAACTCC AGGATCCATA AAACACCCCA AGTTCTGTGG
9651 AGACCCTATC ACATCTGCAA AACTCTCCAG GAAGTCCAGA GCCCTCCTGG
9701 TTAATTTGTT TTAGGGACTA GGCATGCGGT ATCCCCTGAC AACACTGGAT
9751 CAGCAATTCT CCTACCTAAG TCAGTCCCAC ACCATGTGCA GCAGAGTATC
9801 CAGTGCCCCT GCCCTGGTCT GCTCACATTG GTTTGCTCTC CAGAATAATA
9851 ATTCCTCAAT ATCCACAAGA GATTGATTCC AGAACTACTC CGAGGATACC
9901 AAAAATCCTC AGATGCTCAA GTACCTGGTA TAAAATGGCA CAGTATTTGG
9951 CATATGACCT AGGCATATTC TCTCCCATAT ACTTTATTTA TTTATTTATT
10001 TCGGGACAGA ATCTCATTCT GTCGCCCAGG CTGTCACTCG CTTATTGCAA
10051 CCTCTGCCTC CCAGGTTCAA GCAATTCTCC TGCCTCAGCC TCCTAAGTAG
10101 CTGGGACTAC AGACGCATGT CACCACGCCT GGCTACTTTT TGTATTTTTA
10151 GTAGAGACAG AGTTTCACCA TGTTGGCCAG GCTGGTCTCA AACACCTGAC
10201 CTCAAGTGAT CCGCCCACCT TGGCCTCCCA AAAAGCTGGG ATTACAGGCG
10251 TGAGCTACCA CGTCCAGCCC CCCATATACT TTAAATCATC TCTAGATTAC
10301 TTATAATACC TAATACAATG TAAATGTTAT ATAGTTGTTT TAATGTATTG
10351 CTTTTTTTAT TTGTATTGTT TTTTATTGCT GTATTATCCT TTTTTATGTT
10401 TTATTTTTTC AAATATTTTC TACCCGTGGC ACCCACAGTT GGTTGGTGGA
10451 ACCTGCGGTT GGTGGAGCCC ATGGATGTGA AGGGCTGATA GTATGAGAAA
10501 ACTCAGAGGT GCAGAGTTGG AGAGCACATC GGGGAGAATG TCAGCATGGG
10551 TTAAAAAAGA CACACTGTGG TTGGAGATGA TCACATGAAT GGCCACTTCA
10601 AAAATGAATG GGTCTCATCC TCAAAGCAGG CTCTCCTGGG CACTGCTTGG
10651 GAAGGTGCTA ATTGGAGCTT CAGGCAACAA TAATAAGGGG ATACAGGTGG
10701 GGATCCTGCC ATGGGCGTAG CTTACTTTCT CTGGACTCTT CTGGGTCTTA
10751 AGGCCAGTTT CCTCATCCAC TCAAAAGAAT GACAGCAAGG TGAGCAAAGC
10801 AAGGCAGGTA AATGAGGAGG ACTCTTTCTG GCTGTCCAAC TTTTCATCAA
10851 CTTCCCAAAG GTTTTTGGAT GGGACATGAG CACTCATTCC TTCTCCACCC
10901 TTTAGCTAGG CCCTGTCAAC TCCAGGAGGA AGGTAGAAGA GGTCAGAGCT
10951 GTGGTCTTTC ACTTATTCAA GATGTTTCCT TAGTGTTTTG TGTTTGGGTT
11001 TTTTTTGTTT TTTTTTTTTT GACAGAGTCT TGCTCTGTTG CCCAGGCTGG
11051 AGTGAAGTGA AGTGGCATAA TCTGAGCTCA CTGCAACCTC TGCTTTCGAG
11101 TTCAAGCGAT TCTCATGCCT CAGCCTCCTG CATAGCTGGG ACTACAGGCA
11151 TATGCTACCA TGCCTGGCTA ATTTTTGTAT TTTTAGTAGA GACGGGGTTT
11201 TGCCATGTTG GCCAGGCTGA TCTCAAACTC CTGACTTCAG GTGATCCAGC
11251 CACCTTGGCC TCCCAAAGTG CTGGGATTAC AGGTATGAAC CACTGCACCT
11301 GGCCCCTTAT TGTTGGTTTT TAAAAGAGAA ACTAAGCTGT GCTTCCAGAA
11351 CCCAGTTTGA GAAAGTTTGA AGACCTGGCA TAGAGCCAGT GACATATAAT
11401 TGTTAGTTGA AGAAAGAGAG CTCCTTGATC TGCAAATAGA GCACGGCCCC
11451 ATATTTAAAT TCTGCACATT CTAGAAGCAT TTTGCAAGAA TCAAATGCTT
11501 TGAGGATTTT GCTAAATAAC CATGGAGGAA AGCACTAGAC AAATATTTTC

FIGURE 3, page 3 of 57

11551 AGATGGCATG AGAGTTATCA TTCATAGGAA TTATATTTCC ACTCCTACCA
11601 CTTACTGGGG ACCCAAGTAA GAAATTACTT GGATAAGCAG AGGAGAATTT
11651 AAAGTTGAAT GTGGTGGAAC TTATTATGGA AAAAATATGT TTTTCTGAAA
11701 ACTGGATATG TGTATATATA TAAGTTCAGT TGTCATTTTG GAACCATCCT
11751 TACTCTTCCT AGCTAAGGAT TAGCATACAT AGGTGCAACT TGACTAACTC
11801 TGCCTGGACC CAATTCAGTT ACCTTTTGGT GGGTAGGGTT CATGAAGAAG
11851 CAGTTATTTG TGGAGTGTAT AGAAACCACT CTATTGTAGG TTCTTTAGTT
11901 GGTACTTTCA AAATAAGTGA CATCCAAATA GTAACTTAAT ATTCCAAATA
11951 TGGCTGCAAA ACAAATTGTC GATTATGGAT GACTACTACT GCCATCTCTC
12001 CATACCAGTC CATCTTCTGC CAGGCTGTTT GGTCTTGATT TGTCGACCTT
12051 TTAGGTTTCT CCCCATGTAT TCCACATGAC CTTCACCAAC CCCACTTCTA
12101 TCTCCAAACG TCTTTCTGAG TTGTGGGGAT GCAGATGTAT TCTGCCACCA
12151 TCACAAGGGC TAACCGAGCC CTGGCTGCGG ATCTTCATTG TTGTTCACAT
12201 TATTTCCATT CTTACACCCT ACTTCATGTT TGTACACTAT TTTCTTACAT
12251 TTGCTGTCTC TTCTAAACAT TCTTTGCTGC ATCCACTTTT TCTCTATTTG
12301 TGCTCTAGGT GCTGCAGAGG CTAATGCTGG GTTTCCTTTC ATTCCTCCTT
12351 GCACTCAGCA CCTCCCTTCT CAATTCCTTT TGCCATGTCT CCACTTTAAA
12401 TCTTAACCTA CTCCAGATAG TCTTTTCCTT CACACTATTG GCATCTGTGC
12451 TTGGGTTGCT TTCAGTCTAT TCTCTGATCT ATGATTTCTT TGCATGATCA
12501 AGAAGGTGCC ATGAAAGGAT CCCTTAAGAA AGCCTGTCAT TTAGCCAGAA
12551 CGAACTAGCT TCATGATAGC ACCAGGAAGA CTGATATCTC CCAGGAAACA
12601 AACCACTCAT GGTGGTGCTC TTTTTGCCTT CACTATGAAG TGTTTGTCTG
12651 CCTGTATGTG AAAACGAGAG GGTTTAATTG TAAGGATGCA GCACAGATTG
12701 GGACTGGCAT CAGAAAGCCA TTGGGGACTG AGGTAGCTCT AGAGACCGCT
12751 TTCTGTCTCC AGTGCTCTCC CTCCTGGGTG ACATGTTTTC TGTCTCCTGG
12801 CATCTCTGCT TCTCTCTATG GGCTTCTTTA TTATTTGCAG CTTGCAATGG
12851 TACCCCAAAG TCCTAGCTCA TGGCTCCTCT CTGCATATAT GCTTTCTGTT
12901 CCTACCCACA AAGCTCTTTC TATTCTTCTA GTTTAAATTT TCAAGAGAAG
12951 AAATCTGATT TTTTTTTAAC CTGGTCATGT CAAAGACCAC TGACCACATA
13001 TGAGCTGGTT GCCCTGTGTC AAGTGCCCCC TTCTCCCACC CTCTTCCCCT
13051 CCCCATCTGG TCTGTCATAA CTGAATGATG GAGTGGGAAA TTGAAATTGC
13101 CATGGGAATT CCATGATAAG CTATCTAAAC AGTTTTATCT ATAAGTGGTA
13151 GACAGAGTCA CTTAGAAGGG AGTCCCAGGT GAGACAGGCA CCTGTCAACT
13201 CCAAACTGGC ACACATTCTA AGGTCTGCAA CACCCCAGAG AGAGCACTGA
13251 TTTTGTAGTG GCCTGTACTG GGGCGGTAGG CTGGAGAATG GGAGAAATAG
13301 CCACTTCAGA ATCCCCCAGC CCAAATGCAT CAAGCTCACT ATAGACTCTG
13351 CAGCCACGAT TCAGCTGGCT TCTGCTCAGA TCAACAGAAA ACATTCTTAG
13401 TGAATGATGC TTGTGGCACA TATCTCAAGG CTACCAGGGT CATTTCTTCC
13451 CATTTACTTT TTCTCTGATC TATCCTCTCC AGGACACTAG CGTCAGAAGA
13501 TAATCTTCCG TCGTTTTCAG GTACACTATT TGGGTACTGA GTCACTTTCA
13551 AAGCCTCTTT CTGGGTTTGG ATTTCCAGAG CAGCCTGTGC TGTAAAGCAA
13601 GACAGAAAGC TTCCCTGCCA TTCATGCCTG CCAGGGATAG AATGACAGTA
13651 CTCCTGAGGC TCTCCCTCCC CACCCCTCCC CTGCTGGACA GCTGATCTGC
13701 TGGACTCAGC CAGAGCCAGC AGGCACCCCC TCTTTATCCT AGGAGCTGCA
13751 AACTTGATGC CTTTCCAGGA AATCCCCAGA AGCTGGAGTA TCCTCATCTA
13801 CATGTGGCAC AGTGTATGGT TGTGTCAGGT GCTCATGTCC CATTGCATAG
13851 GACTGGGGTG GAAAATAGGG ACCGTCCTTT TGTGTCAGCT CCAGTCAATG
13901 AGTAGTGGCC ATCCAGGGGG CCATCTTGGA AAGGACTTGT GAGGCTGTAT
13951 CTGCGCTCAG TTGTAGATGT GAGAAGAAAA GGCCAAATAT CTGCCAATCC
14001 TAGTCCTGGG ATTCAAGATA GAAAGAACTG CATGGAGTGA AGAAACTAGG
14051 AGTCTCCATT TCACTGAGAT GCATAAGAAT GAAATTATTG TCACTATTTC
14101 TTCAATACTG GCCAATCCT AATAAGAAAA CCCTTTTTGA GTCTCTCTTT
14151 TCTTTATCCT ACATATAACA CAGAAGCTTT TTCTATTCCC TGGATGAACC
14201 CACAGGGACA GAAATTCTTG TTGGACAGGT GAAGCAGATA ATTTCTTTAT
14251 CAGACTAGAA TCTTCCAGAA GCACTGCTAA CCTAGTGAGT TTTGTACTCT
14301 AGACAGGTGG TTCTCAAGCC AGCTCCCCAC CGCAGGCCTT TTTCATGGTC
14351 TGCCCCTCCC TGTGGAACCC ATGTTTTAGG TTATTAGCTG ATAATTGGAT
14401 TTCTATTTTT TCTCATAAAA TACAGCAAAA GATAGCTAGT GATATTATGA
14451 TGAGTTAATG TAATTATAGC CAAAGCAGAG AGAAACAACA TTTTAATTAA
14501 CCTGTGTGGA CTGCTGGAAG AATATAAACT TTCTATTTTG GGGGTTGAGT
14551 AGAGACAGAA ATGAACACAG CCAAGGGCTG ACTGTCAGAG GACATTTAAC
14601 TGATGTAAAA TGCTTTGAAA TTATTGGGCA CTCATTGTTT AAAGTTGTTT
14651 TTGATGATGG TAACTCCGTA AGGGGATCAG AACATGCTGG AAAGAATGGG
14701 CACAGCTTTG GTTACCTGGG CCTTACCACT GTTATTCAGG CCTCTGAGAA
14751 AGCTTACTAT TGTTGTTATG TTTCTTACAT AATAAAACTT CTAATATTTG
14801 TATGAAAACA TAGAATTCCA CTTTTAAAGA TGTAAGGATT TTGTCATACC
14851 ATTAGGGTTA CTATGATCAC TTGATTCTAG GTCTAAGAAA TATTAAGTAA
14901 TTTACCCGCG AACACAGAGT TTTAAGGGTA AGTATCAAAA CCTTGATCTT
14951 CTAATACCAC ATATTCTCAC TCATATGTGG GAGCTAAAAA TATTGAGCTC
15001 AAAAAGGTAG AGAGTAGAAT TGTAGTTATT AGAGGATGCG AAGGAGGATA
15051 GGGAGAGGTT GGTTAATGGA TACAATGTGA AGTTATGTAA GAGGAGTAAG
15101 TTCTAGTGTT TTGTAGCACT GTAGGGTGAA TATGGTTAAC AGTAATTTAG
15151 TGTATATTTA AAAAAAAATA GACAGGATTC TGAATATTCA CAAAGAAATG
15201 ATAAATATTC AGCTGGGCGT GGTTGCTCAC GCCTATATTC CCAGCATTTT
15251 GGGAGGCCGA GGTGGATGGA TCACCTAAGG TCAGGAGTTT GAGATCAGCC
15301 TGGACAACAT GGTGAAACCC CGTCTCTACT ATAAATACAA AAAATTAGCT
15351 GGGCATGGTG GCGCACACCT GTAGTCCTAG CTACTTAAGA GGCTGAGGCA

15401 GGAGAATCGC TTGAACCTGG GAGGCAGAGG TTGCAGTGAG CCGAGATCAC
15451 GCCACTGCAC TCCAGCCTGG GTGACAGAGT GACACTCTGT CTCAAAAAAA
15501 AAAAAAAAAA GAATGATAAA TATTTAAGGT GATAGATATG CTAATTACCC
15551 TGATTTGATC ATTACACTTT GTATACATGT GTCAAAATAT CACTCTGTAT
15601 CCATACATAT GTATAATTAT TATGTGTCAA CTAAAAATAA AAGGAAAAAA
15651 ATCATTTCAG TGTATTTACA AAACATATGT AACCATTAAG AATAATGTTT
15701 TAAATTATAT CTAAGGGTGT GATAAAATTA CAGTATAAGA TTGTGCTTGA
15751 AAAAGTGCAA TAAGAAGTAA ATATGTACAG ATGAGAAAAA GTGCAAAGAA
15801 CTAAGTCCTA AGCAGACTAT ACCTTTCCTA CTGCATGGTA CTTCTCTGGC
15851 CTTTTGCTTT GAAAGATTTT GCACCCAGCA TGGCAAGTGG TTAGCAGAGG
15901 CAGCCATTCT CACTTGTGCG TTGGCTTTGG GAGCCATATA TGTTGTTCAG
15951 CTGGGTGTGG AGTGGAAAGG CTGCATGTTG TATTAATGCA TTGTTAAGAA
16001 CCTCTAAGAG TGATTTCTTT TGGGAAGTGA GACTGACGGT CCGAATGGTG
16051 GAAAGACAAC TTTTAATCTT TTACTTTACA CTTTGTGCAC TTTTAAATGT
16101 TTAACATGAG CATGCATTTC TTTAATAATA AAAATACAAA AAAATTTTAG
16151 CCCTAGATCT TCTGATTTTA AACTGCATAT TCTTTCTATT GTGTTACATA
16201 TTTTAGCATG AGAATAAGGT TATGAAGCTG GAAGTAGCAG GCTCCCTTTT
16251 CCTCATATGT AGGAAGTTAA GAATGCATTC TACGTTTCTT CTTTAAGGAG
16301 TTGGCTTCTT TCCTTTTAAC ATAGGGGTAA CTGGGCCCAG GGAGTTTGGC
16351 AAGGGCCAAA TAAAGTCCTT AATGCCCAGC TCAGAAATCT GGATTCACCA
16401 TCCTTGACTG CTGGCTCCAA CCCACCCTCA CCTGAGCTGG TCTGCAGAGG
16451 ATTCTTGTTT GTGTCACTTC ATCACCAGCA ACTACCGACA GATGATGCTT
16501 TGGCCTGCTG CCTGGGTAAC AGGGCGAGGC TGGCTCAGGA CCATGTTTTC
16551 AGATCAGGGG ACCTCCTTTG ATGCCATGTC CATGGTGTCC GAGGGCAGCC
16601 AGGATCAAGG GCTAGACGGG GCAGTGATGA GATGAGAGCA GGAGGGGCTC
16651 AGCTGCAGCC CCAGGAGAGC CTATGCCAGC CCTGTTGACC AAGGAGGACA
16701 GAAGCAACAG GAGAGCGGAG GCAGAGGGGT GAGTGTCTAT CGCTCAATGT
16751 ATAATCGGCA GACATTTGGG GAGCTCATAC TGTGGGCTAA GCACAGGGAA
16801 GAAAGGCACA GTCCCTGTCC TCAGGGAGGT CACAGTTGAT AGGGAAGACA
16851 AGCATATGTG CTAGCTGCTA TAGAAGGGGG AACCACTGAG GGCTGTGGCC
16901 ACACAGAGGC AACACCCCCT TCTTGTTTTT TTGTCAGGGA TTCAGTTTGG
16951 CGTCATTAGA AGTGACTTGC ACAACCCCCT CCTCCAGTCA ATTCAGAAGG
17001 ACTTGTTAAG CAGGAATGAT GAATTAGCTT CAGCTTGTGG GGCACACACA
17051 GATGGAAGTA TAAGGTGGCC TCAGGAGTAA GTAAATCCCC ATGCAAGCTG
17101 TGTCCTTAGA CCAGAGCAGC ACCCGGTTCT TCCCCATTTC TAGTAAAGGT
17151 GCCTCACACA CCACCAGGAC ACAATTTATG CCTGCAGAAT GAATGAATGA
17201 ATGAATGAGT GAATTCCTGG AACCTCTTCT GCTTATGTGC CACACCAGGT
17251 TGCAGCAAGC CCAGGGACAC CTGGGACTGG AATTGGGCTC TCAGGTGTAA
17301 GGACCAGGGA GCACCCACCA TTTTGCATTC TTCAGCCCTT CCTCCTCTCC
17351 TGTCCCAGCT TCAGCAATAT CCACAGAGCC CTCTGAGCAA CTCTGAGCCT
17401 CTCCACAGCC TGACGCCTGC CTGGGCACCA GCTCTTCAGA GGGTGTTTCT
17451 GTGCTGCTCA GCTACCTCTG AGCCTGGGCT GCCTTTGATG CTCAGGAGAC
17501 ACCCTGTAAT TCAATTAAGC CTTCTCTCCA GGGAGCATGT AATTATGTCC
17551 TATCTGGGCC TTGTAATGAC AGCCCCCTGC CACTCTACAG GGAGTTGCCC
17601 TGCTCAGCTG CCCAGAACCT TTCCCTGGGA GGAAACTAAT CTGCTTAGCC
17651 CAGATTGGAC GCAGTTCTGC ACAGCACTTT TCCGAATGCC TCTGAAATGA
17701 GTCCTCACTG ACAGAACGGG CCCACTCTGG GGGAACTGAG GGCTCTCTTG
17751 GTCCTGCACT GCTCTTTGCC ATACAGATCT GTCTGCCCAG GATTTTTCTT
17801 GGGTGTGTAG GAGGCTGAGA GAGCTCCCCT TTCTTCTCAT GGCTAAATCC
17851 CTTGGTCTTT CCAGCCCTCC TGGGGGTTAG AAGGGAGAGG GAAAAAAAAA
17901 AAGACTGAAC TTGTTGTTGT TGTTTTTGTT GTTGTTGTTG TTTGCCTGTT
17951 TTCTATGTTG TCTTGTGGGG AGAGGGTATA AGATTGATTG ACAGAGTGGC
18001 ACACTTCCCC TGCAAATTCA TCATTTGAAT TTCTCAGGTA AGATGTTCAC
18051 ATTTCTCTGT TAAGATGCTC CAATTTCTCT GGTTAAGATT TCTCTGGTAA
18101 GATGCTCATG AATTGGTGGA GGTGTTGGCG GGATGTGGGA AGTGTGCCTG
18151 CTCTTTCTGA GTTTTGGGGG AAGTTGCCTT AATTCTCTGC ATGACTTTCT
18201 TTGCTCCTTT GGGCTTCATT TCTGTGCAAT GTAGTCTGAC ATGAATACTG
18251 CTCAGGGAGG TGTTGCTTCC CACTGCCCAC GCCACTGGAA ACCAGTAGCC
18301 CAGGTTTACT CGAGTCCTCC TTTTGAGGAA CCCAAATTCT TTCATTTCTT
18351 TTATGTGAGA TCTGCCCAAA ATGCCATTGG CAAGCTGTAC TGGGTTGAAT
18401 AGTGTCCTTC CTCCTCCCAA ATGTATGTCT ACTCCAAACC ACAGGATACT
18451 ACCTTATTTG GGAATAGGGC TTTTGCAGGT GTAACCATTA ATAGTTATGA
18501 TGAGGTTATA CTAGATTAGA ATGGGCCCTA GATCCTATGA CTGGTATCCT
18551 TACAAGAAGG CCATGTGATG ACAAAGACAA AGAATGGAGT GAGGCACCCA
18601 AGGAACTCCA AGGATTGCTA GGAAACACCA GAAGCTTGGA GGAAGGCATG
18651 GAACAGATTC TCCTCTCGGA CCTCTAGAAG GAATCAGTCC TGCTGATACC
18701 TTGATTTTGG ACTTCTAGCC TCCAGACCTG TTGGGGAGAA TACATTTCTA
18751 CTGTTTTAAG CTACCACGTT TGTGGCGATT TGTCACAGCA GCCATAGGAA
18801 ACTAATACAT ACAACCTGCA CAATGCCTAC TCCAGCATTC CATAGCAAGT
18851 CAAGGGCCTC ACAATTATGT CCAAAGGACT GATAGAAGAG CGACCTCTGT
18901 GCTACTTGTC CCTCAGGACG CTGACCCACA GCTCTCAAGG CAGGAGTAGG
18951 CCAGAGCTCA TTCAACAACT TTGTTATATA GGGGTTCCAA TTGTAAACCT
19001 TTTGAATTCC TGTTTGCAAG TAGATGAGGG TTGAAAAATA AATGGCCACT
19051 TTCTCTAAGC CACATACCCC AATCTGTTTT GTTACTTCAT TACAGCTGTT
19101 ATAATGGCCT CCTCTTCTAT CTTCCAATCT CCATAGCCCT GGTTCCTTGA
19151 TAGTTCTTTT TTTTTTTTTT TCTTTTTTTG AGGCGGAGTC TCGCACTGTC
19201 GCCTGGGCTG GAGTGCAGTG GCACGATCTC GGCTCACTGC CACCTCTGCC

FIGURE 3, page 5 of 57

19251 TCCCAGGTTC AAGCAAGTCT CCTGCCTCAG CCACCTGAGT AGCTGGGATT
19301 ACAGGCACCT GCCACCATGC CTGGCCAATT TTTTGTACTT TTAGCAGAGG
19351 TGGGGTTTCA CCATGTTGGC CAGGCTGGTC TTGAACTCCT GACCTCGTGA
19401 TCCACCCACC TCAGCCTCTC AAAGTGCGGG GATTACAGGC ATGAGCTACC
19451 GCGCCTGGCC AGATAGTTCT TAAACAACTG CCCAGAAGTT CCAGCCTAGG
19501 CAGGGGCAGC CATGAACTGC ATTGCTCATT TCTGCTTTTT GACCTTTTCG
19551 ATGGCTGAAC TCTAGGCCAT GGAAAACAAG GACCCACTGT ATAGTTAAGA
19601 GTCATTTTGT GACTAGGGAG ACAAAAAAGG GCCTATTCTC CAAATCCCCT
19651 TTCCCTCTGG AGTTCCTCGG TGCCTTAAAG CTTGTCCTGA GCTACAGGTG
19701 TGTTACCTGC TTATCCCAAA ATGCAGGCAT GTTACCTGCT TTCCTCTGCA
19751 AAGAGAGGCA GGCCTGGCTG GGGCACAGCT GAAGATGTCA AGGCCAACCT
19801 AAGGGCAGCC AAGCTATGGC TGTCTGTGAC AAGAGGAGAG CAGCGGTGAT
19851 GGGAGGGTAG GAGGCATTGA GTTCATGTCC GGGTTTGCCT CCTACCCTCC
19901 TATCACTGCT TGATGATCCT ATCACTGTCT TGATGAGTTC AAGACAGAAG
19951 TTTGCCTCAT CATTGCCACA ATAAAATCAC CAATAACAGA AGTGTGAAAG
20001 CAGCGATGTG AGTGGAAGCC CATATATACA CAGGGGGTAA TAGAGCAGCA
20051 TGATTAAATA TGTGGCCTTG TTATCAGACA GGCTGATTTG GAGTCCCAGC
20101 TACTTGTTGG TGACCTGAAC TAGAGGAAGT TATCTAACCT TTCATTTTAC
20151 TCATTTACAT AACATGGCTA ATAATAGCAC CTACCTTATA GGGTTATTGT
20201 GAGGATTGAA TACAATTATG CAATATAAAA CGTTTAGCAT AGTGCCTAGT
20251 CTAAATTCCT CACCAGGGGT ATGATGTACT AGTTTTTAGT TAAGTAATTA
20301 GTATCCTGGA CATGTCACAG CCATTTGACC TATCTGGGCC AGCGTTTTGC
20351 TCAGGTTCCC CCAGCAGTAA TTGTATTCCC TCCCCAATCC CGGGATTAGC
20401 TTTTAGGAAG AAACAGTTGA TCTAAAGATA GAAAGTCAGA GTACTGTCTG
20451 GAGGAAGGTA GAGGGAAATG TCATTATCTG GGTTTTCTTT GATGATGTCA
20501 GGGAACATGA CAGGCTGCTC CCAAAGACAG AGCAGCCCCA GGACAGGGAA
20551 GAAGGTGACC TTGAGGTTGA CTCCTCTGCA TCCCGATGTG GACGTTATGG
20601 ACTTGTTTTG GAGATGAAGG GAAAGAAAGA TGGAATGTAG AAAGTGAAGG
20651 AGAATAAAAG AAGTGGGAGG AAGAAGGGCT GGGAGGAGGA TGGGCAAAGT
20701 CTTTCTGGTC TCAAGGATAA TTACATGTGA AATCACTTGC CAGTGGGACT
20751 CTGGGGCTGG AGCAGCTACA ATAATTACAG TACAGGCTGC AGAGGGCTCT
20801 TGGGCATGTC TTGGAGCAGC CTGTAGGCAG TACTGAGGCC TCTCTCACTA
20851 GACCCATCTC CCAGATCACA TAGTACACAC ACCTTCCACC CCCGGGCCTG
20901 TTAATGATCA AAAAGCTTAA ACAGAACAAT TACAGCTTCA GAGTGGAACC
20951 ATATCTCTGG GCTCCTGTGA TGAAAACCAC AAGCCTGTCA GGCTGGGGCT
21001 GCTTCACATG GAGGGCCCTG CTCTTAATGG CCAAGTGATC TGGAGCAAGA
21051 CCCGTGACTC TCCCATAGTG CTGTGGATGG TGCTGCCTCT CCCCACGCAT
21101 CCCCAGAAGA GGAAGTTCAG TAACTAAGGA ATTAACTATT CTCCAGCCTG
21151 ATTCTGCTTT TCCCAATCAG GGCTTTATAC CTTTCTTTTT CATCCCTATA
21201 TTTGGAGATG AGTCACCCTT GCCTTCATTT TACCTAAGCA AGGCAGTTTC
21251 CTGTAACCTA ATGAAGTGCC AAACAATACT GTGATTTATT TAGTACTTAC
21301 TGTGTGCCAG GAATTCCAGC AGGTGTTGGA CATTTATGAT GTATGATCCT
21351 TACACTAAGC CTGCAATGGT GCAACCCCAG CCCTGACCAC TCTGTGCTTC
21401 CCTTTTCACA ACACAGCTTG TCACTAAATC CAAGTCAGGA ATTCCAGGTT
21451 AGGCTTGAGT TGTGCAGAGC CCTTAACTGA AATTTGCCAT GGTTGAGGCA
21501 TGATTGCAAT CACTGACAAC TCCTCCCGGC TCTACACACC TACTTGTCAT
21551 ATTCACGCCC TGATCACGGC CCCACTCGCA TCTCTTCCCA CTTTAGAAGT
21601 TCTTTCCTAT AGAACACGTT GCTGCTGCCC TGTTCTGGTC ACTGATCAGC
21651 CCTGGCCTAA CCACTGGCTA AGCTTTGTGC TTGCACATAG CTGGTTGAAT
21701 CGTATGTATT GCTGTTTGTG TACATCAAAA ATATAATAAT AATATCGGCA
21751 ATTTTATGTG TTTCATTCAA CATGAGGGAC CCAGCATTCT TACCTTGTCG
21801 CTTTGTAAAC CCTGCTGCTC TCAAATCTCC ACTAGCTGTT TCCTGAGCAG
21851 AAGGAGATAA AAGGCTGGCT CACACCCCCA TGTTTTTACT GGTCACAGTT
21901 ACTGCCACCA TCCAAGGCTG AAGAGACTTC CTTTGTGTTA GGGCTAAAAC
21951 CTTAGTCATT GTATCTAAAT GTCTTCTGTA TTCCTTTCCT CAAAAGAAAA
22001 AAGTACCCTC TTCTGCCAAC CCTCTCCCAT GCCAACTAAA CAAGCAAGCA
22051 AGCAAACAAC AAAGAAAAGG TGATATTACA GATGCTGCTC AGCCTATGAT
22101 GGGGTTACAT CCTGATAAAC CCATCACAAG GGATGTAATT CCATTGCAAG
22151 TTACAAATAC CATAAGTCAA AAATGTATTT ATTTCATATA ACCCACAGAA
22201 CGTGATAGCT TAGCTTAGCC TACTTGATCA TGTTCAGAAG ACTTATATTC
22251 GTCTACAAGT GGACAAAAAC ATATAAAACA AAGCCTATTT TAAAATAAGG
22301 TGTTGAATAT CTCATATAAT TTATTGAATA TTGTACTGAA AGTGAAAAAT
22351 AGAATGGTTT TCTGGATACT CAAAGTATAG TTTCTACTGA ATGCATATCA
22401 CTTTTGCACC ATCATAAACT TCAAAAATTG TCGGTCGAAC CTTCCTGAGT
22451 CAGGAATCCT GTCTGTACAG GGTATAAAGG AGGAAAGCAT CAGCTTTGGA
22501 GGCAGGTGGA CCTGTGTTTG AACCCTGATT CTGCTAGAGC TTGACAATGC
22551 ATATTCGTTT TCTATTGCAT AACTAATTAC TACAAACAAC ACATTTATTT
22601 CTCAGTTTTC ATGAATCATG AGTCCAGGCA CAATTTAGCT GCAGTTAAGG
22651 TGTTAGCTGG GGCTGCTGTC TTATCTGAAG CATGGGGGTG GGGGTGTGGA
22701 TTCCAAGGTC AGGTGGTTGT TGGCAAAATT AATTTTCTTG CAGCTATAGA
22751 ACTCATGGCT TGCTTCTTCA AGGACACGGG GAGAGAGAAT CTCTCACATC
22801 TTTTAAAGGG TTCACCTGAT TAGGTCAGGT CCACTCAGGA CAGTTTCCCT
22851 TAAAGTCAAG GCTTAATAGT CAACTGATTA GGGACCCTAA TTATATCTGC
22901 AAAATACCTT CACCATTGCC ATGTAACATA ATCATGGCAA ATAATCACAG
22951 GTCCCAAATG TTCACAGGTC CCACTCACAC TTGAGGGAGG GGATTATATA
23001 GGGCATGTTC TTGCGGAGAG AAGGAATCTT ACAGCCACAT TGGAATCTGT
23051 CTTCCATGCT ATTTGACCTC AGGCAAATTG ACTAATCTCT TGAAGGTTCA

FIGURE 3, page 6 of 57

23101 ATTTCCTTAC CTGGAATAAA AGGACAATAA GATCAGCCAT ATAAGGCTAT
23151 GACAAAGACT AAATGAGATA GAATAGGCTG GAAAAGTCTT GCAGATAGCA
23201 GACACAAGTA TATAACAATT TCCCTCCTAC TGTTCCTTTT GTTTTTCACC
23251 TATCCTGCAG TCTCTGTCAC TTCAAATACC ATAGAAAACC TTTCCAAGCA
23301 GCCCAAATCA TGCCCCCAAA TAGTCACGTC TCATTATTCA TAGCAGTTAT
23351 GTTCCATAAA GTTAGCACAA ACTCCGAATG AGTGAATCCT AAAGCGTTGC
23401 TCCTGGAGGA AATACAGGCT GCTGGTCACA ATATTTTTAT CAACTGATCA
23451 ATATATACCT TGTCTTATGT GTGTTTCTGC TTCAAGACAC TTTATTTAAT
23501 ATATACGTTG ATTCATTAAC TCTGAACTCT CTAGGCAACA GCATTATAAC
23551 TCCTGCCTTC ACAAAGCTTA TCTAACACAC ACATTTCCTC CTCAGGCACA
23601 TCCCAGCCTT CTTGCACTTA GGATTCAGCA GTATGCTTAA GGGCCATTTT
23651 CAACAGCAAA CTCATCAGCG CAAACACAAA CATGTGAAAA ACGTAGCACT
23701 AAAGAGACTG CAAAAAGGAC ACTGGCTTAC AGCATGGAAG CTGGAAGGAG
23751 AAGGCAGAGA ATCACCTTGT TCCACTTCAG CTATGAATAT GCAGTCAGGC
23801 CACCCAGTCA TTCAAATTTT ATAAATATAC TCTAATATAT ATATAAATAC
23851 CAGGCAGGGT TATTTTTTTC CTCAAGTCAT TTTTCTAATT TTTTTTAAAT
23901 GAATAGATAG AAGAGCTGAA GTAAGGGTCA GGAGCAAGAG CTCTGCTTCC
23951 TTTTCCCTTG CTGGGCTTCG TTAGAGAGCC ATCATCTCCT CAATATGTCT
24001 CCCAACTCTT CTAGGCATTG GATGAGTTTG CTGCAGATAC GAAACCCAAC
24051 TTTGCCAGTC ACTTCATACT AACAGGTGAA ATGTAGTGGA GGAGCCTTTT
24101 GAAGACAGGG ACTCAGCCCC CCATTAGCCT CATTGCAGAC CTAGATTCCT
24151 GCCAAAATTA ATTTGGCTGG AACTTCCCAG CCATGGCATT GTCGACATTA
24201 CACATCTTCC ACTGTAATGT CAATTACCAT TTTATTCAGC CGAATGCTGG
24251 AGAGTTAATG TTCAAGTGGT TAGAGCTGGC TACGGGTGGG CTGAACAAGA
24301 TGTCTTTTCC TTCATTTCCC CTGCCTGTGG TGAAGGATTG TAACCAGCCC
24351 TGGCTGGCAG CACTTTGAAG CTCACCCAGA GTGCTCCTGG GGACATCTTC
24401 TACAGAGCCT ATCATTTGGA CATGCTGTCT TCTGGGCCTG TCTTCCTTCC
24451 TTCCTTCTTC CCTCCCTCCC TCCCTCTTTT CCTTCCTTCC TTCCTTCCTT
24501 CTTTCCTTCC ATCTGCTTTA AAACCAGCTG CCTTGAGTGC TTGTCTTGGC
24551 GCCCCTCATT AGTGCCATTG CAATCATCCC TCCTGCCTAC CCTGCTAACC
24601 ACAGCTTGTT AGTCCACAAC AGCAACAGCT GTGTGCTGGG GTGCAGCAGC
24651 TGGAGGGCCA AAGGTAGGGC TGGGGGACAG GGTGTTGGGA TGGTTTTCTG
24701 GGGCAGATGA GTTTATACGT TTCTTTCATG TCCCCTTCCT CCCACATAGA
24751 CTTTTATTTC CCCAAAGGAA AACAGAAAAC AATGATCTGT TTGACAGTGT
24801 TGCTATCATT GGGCATCAAA CCTATCATCT AAGGGGAATC CCCCTGTATA
24851 ATCAGTCAGC CAAATGGAGC AGGACCCTGT GTTTTGTAGC TGATACAACA
24901 GGGCAGCATC TCTAGTGAGG GGGCCAGGGC TTCTATTTCC TTCATTAAAA
24951 AATGAAACAG CAGACCTGAT TCCATATTTA GAGATTACAC TTAGTTGCCA
25001 CTGTGGGTGT GCAGGCACCA ACCAAACCCA GTTGGCACCG TTGTCTTTTC
25051 TCTGCAATGA TGTATTGAAT TTAATAATGG AGGTATATGA AATTCAGAGT
25101 GATTGGAACT GAAGGTTTAG GGGCTTTGTG TAAAATTGAT ATGTAAGGGA
25151 TTTGGAAGTA GGTGAGGGAT TCTTCCCCAA TACTTATTCA ATTTTGGAGT
25201 CAAATAACCA AGCATTTACA AATAGCCAAA AAAGAAATTG AAAGAGGGTT
25251 TAATCCAATA AATTTTCATG CCTCATATGA ACCACATCTT ATAATAAGAA
25301 TTATGCTTTT TCATTTCATA CTCAGTTAAC AAATATGATT TGTGAGCACC
25351 TGGTAAGTTC AGGGCACTAG GCTGAAAGGG GTTACCAAAT GTCTTCATTT
25401 AACAAAGTCC AGCTGAGCTC TTACAGGTAC CAGAACTGTG CCTGGGCTGT
25451 CATATGAAGA TGAATGTAAG AGTGTGTCAG GCCTTCAAGA GCTTACAGTG
25501 TGTCAGGAGA CATCAAACAA GTGAGCCAAT AAAATGATAC TGCCATTTTA
25551 GAAATAGCCT GAAATTCATG GAGTTCACAG TCTTGTTAGG AAAGTGAAAC
25601 ATAAACCTAT AAGCATTAAA AAATAACTGT TGAAGACAGT AACGGAAGAA
25651 TGCAACTGGC AACTGAATGA TATAGGTTGT GATGACTGTT AAATATCATG
25701 AAAAGAGACC ATGATGAGCT GAGGCACTCC AAGAGACTTC TTTTTGGAGA
25751 TATGTTTGGA GCCAAATCTT GAAGATTTAA TTGCTTTTTT CTTTTTTTTT
25801 TTTTTAGGTG GAGTCTCGCT CTGTTGCCCA GGCTGGAAGT GCAGTGGCAT
25851 GATCTCTGCT CATTGCAACC TCTGCCTCCA GGTTCAAGCG ATTCTCCTGC
25901 CTCGGCCTCC TGAGTAGCTG GGATTACAGG CGTGTGCCAC CATACCCAGC
25951 TGATTTTTGT ATTTCTAGTA GAGATGGGGT TTTGCCCTGT TGGCCAAGCT
26001 GGTCTCAAAC TCCTGACCTC AAGTGATCTA CTCGCCTTGG CCTTCCAAAG
26051 TGCTGGGATT ACAGGCATGA GCACTGTGCC TGGCCTTTTT TTTTTTTTTT
26101 TTAAAAAAAA AAAAAAAAAA AACAGGAAGT TTTCGTTAGT TTTTTTGTTT
26151 GTTTTACTTC CCATAAAAAC TCTTTGTGTC ACATGGAGGT GAATGGAAAG
26201 AGAGGCTGTG GCAACAGACG GGAGACTTTT CTGATATCAG AACCCAGTCC
26251 CATAGACCAG AATGTATGCT TTCAATCCAC GTTGTCTGGG TCCATCCTAT
26301 TGAGTGCCCT GCCCCCACAG CGGGGTATGG AGAAGAGTCA GACACAGCCC
26351 CAGTCCTCAC GTAGCTCACA ATCCAGTGGA GGAGACGGAC TCAGAAACAG
26401 ATAGAGATGA AGCCATGAGA TCAGTACTGT CCGAGGCCAT GGCCACGGTT
26451 TTGTGGGAAC CCACGAGAGG GAATGACTAA CTGTGGGGAA GAAGAGGGAG
26501 AGGACCAAAA TGCAGGGGAA GTGCTCACAG AGGATAAGTA AGCAGTGAGG
26551 TGCCATGAAA TGAGTATACA CCTGACAGCC GTGTAACAGC TCAGAGCCTG
26601 GGTAGAGGGG AATAGAGCTG CTGGTTCTCT GGGGGGAAGA GAGGGGTATG
26651 GGATTCTGGA ACAGAAGCAC CAAAACCAGC AGGTTATTGG AGCTGTTAGT
26701 GCTCAGATCA GCAATGGGTG CACAACCAAA CCATTCTCCT AGGGATGAGT
26751 TCTTTCCTGT GGATGAGGGC TTCTCAGCCT GGCTTCTCCC GAGAATTACC
26801 CGGGAAGCTT GAAAAGTACT GATGCCTGGA ACCTACCTCC AGAGAGTTGG
26851 ATTTCATTGT GTTGACGTGG GGCTGGGATA TCAGTATATT GTTTAAGCAC
26901 TCCAGGTGAT TCTGATACGT AGCTGTGATT GAGAACCCTT GCCCTAAGCT

26951 ATCCATCTGC ACTCCAGGGG TGCTCCCAGG CCCATCTGTT TGTAAATGGA
27001 CAGGTGTCTT GAGGTAACAA ATGTGCCAAG GCTCTGGAGC CAAGCACGCC
27051 TGGCTCCTTA GTGCCTACTT AGTGACCTCA GGCAAGTTAC TAAATGGCTT
27101 AAACTTTACA AATCCTTAAT TTGTAAAATG TGGGCAATGA TAGTACCTCC
27151 TCACAGGATT ATTACGAGGT TTACACGGAA TACTCTCAGC TCATAATAAG
27201 CACTTGCACA GGCCTCATGG GCTAGGCCCT CAAAACTTAA CGCATCTACA
27251 GGCAACAGCC ATATGAAAGG AATTTTATAC CACCAAGTCA AAAAATCTGT
27301 GAGCACTGCT CAGAAGCAAA AGCCTGTCTC CAACAGCGCT CATTTAAGGG
27351 GTGGGCGAGC TACAGAGAGA AGAATGAGCC CCCACAGGGT AAGCTGGGGA
27401 AAGCTGGGGA CAGAATGAGA CTCAGGAAAT CACTTGAATA TTGATTATAT
27451 TTGTGCTCAA TAATAAAATA ACGAAATGAG TACAGCCCTA GACCTAAACA
27501 TTGTGGGTGA GGCAAAGGCA ATGCGTTAAT TTTGCATCCA CTGAGGAAAA
27551 ACTCTAAAAC GGTGACTTCT TTTTTAAGGG ACCAGAAGAA TCTAGATTAT
27601 ATTTAGTCTA AGTCAATACA TACGACAGAA CCTTGCCCTC TAGACTTGAT
27651 AAGAAAGAAG TAAAATAAGA GAAAGAATAA AAAACCCTTC CACCAAAATA
27701 CTAACATTCA GATAATGACT TTTTAGTTAG GTCTCCTGGA GAGGAGGTTC
27751 CCTCAGAAAT GAATAGATTT CTCTTCTAGT GCAATCATCA AAAGGTAATG
27801 CATGGACTTA AGTGTGATCC CCAAGAGAAA ATCAATGACC TTTCTGTGTT
27851 TGCCTTTGAG AAAATCAGCC AGTCTATGGT TAAATTAGAC ATATTTTTTC
27901 TCCTTGGTCA AGATTAGTGG GACCAAGAAT GCAGTCTTAC ACTCCTTCTA
27951 GCAAAGAATT ACCTGATGCC TTATTTCACA CAAATTTGCA AAGTTGTATG
28001 GACGTTGTAT CTTATTTTAA GGAGAACTGG TGATCAAATG ATGACTATTT
28051 CAATAGTGGT TCATTTACAC CACCACCCTC ACCCCACATC CTGCTTTCAC
28101 CTGAATCTGA ACGATCATAG TCAGTCTGAG ATTCTGAAGG TTTGAAATTC
28151 CTTTTCTGAG CTCTGCAAGA ACAGCATCTC CCAAGAGAGC TCAGGGCAGA
28201 CTGTCTGGGA GAGATTGGAA ACCTGTCTTT TGCAGTAACA TGAATTGGTT
28251 GAATGGTCAC CCTCCATATC AGGCCTGCTT CTCCCATTGG GTTTCTGATC
28301 AGCCCAACTT GGGTCTCACC CTTCTGATTT CTCTCTCCTG GCTCACATGG
28351 GGCTGCACTG GCCATTAGGT GCCAGGCTTG GCTCCGTGGA ACCCATTGGC
28401 CAGCTGGGCT CTGTGGAGCC CTAAGGCAGG GCTCTGGTCA CTGGTGAGAG
28451 GGAGGCCATT GGAGTCACTG GGGTGGACCT ACAGACCCTA GGGTTAACAG
28501 CTAGGTGGGT GTCCTCTTCA GAGAAACGGG TTACAAAGTG AAAGAAAGTT
28551 ACACTGTGAG GTCAGCCAGG GAGGAAGACA GAGAGCTGAT ATAAGATAGG
28601 TACTGATTCC CTGGGGATGT GAAAGGAGGG TAATATTCCT AAAATGATAG
28651 CATTTAGCTT CCAGTATACA TTAATTGATT CCTGATATTC ATTAAAACTA
28701 AACGCTATTT CCTTGATGTC TCATCCAAAG CCGCACCACT CTTCCCACTA
28751 AGTCTGAGGG GAGCTTGTTT TGTTGACAAG TGTAAGAGGT TGAAGAGGGA
28801 CCCATGAACT CTTTTGTCCT ACTGAAGAGA TCCACAGATG GAAACAAATG
28851 CTCCTACCAC ATTTATGAAC TGCTGCTTTG CAGTCCCGCT TCTGCTATCA
28901 TGCACAGGAA CTGACTAAGC TCCAAAGCCA GAGGATGTAA ATCTCCCTGT
28951 AATAAATGTA AGTCATTTAT TAGCTACATA CACTTCAGCA AGTCACCTAA
29001 CCTGCAAATT TCAAGCATGT GAATCTTGGA TCTTTCATGT GCTAGCTGTG
29051 AGACTTTGAG AAATGTATTT AATGTCTCTT TGCTTCCTTT TCTACCCACA
29101 CAATGGGTAT AATAATGTCT ACCATATATC TTTGCAGCAA GGTCTAAATG
29151 GGGTGATACA TGCTGAATAC ATTTCCAACA GAGTCTGTGC AATGATAAGC
29201 TCTTTCCAAA TGTTAGTTAA AGCTAACCAA CTAACCCACC AACAAACCAA
29251 CCTCTTAGCC AGGACTGATG GAAGGAGTCT GTGAGAGAAT GCATTTAAAA
29301 CACTTGGCAC CATGCCTGAC AAGAGTAAGT ACTCGATAAA TCAGTTATTG
29351 TTATTATCGC ATCGGTATTA TGACCATTAT CCTCTTCTCT ATAGGCTTCA
29401 GGTTTTCCTG TCTTTTTATC ACAGCAGTAT TCCAGCAGAA GCCTTTGATT
29451 TAACTAAGTC TCTACTGTGT GTGTGGCTAG ATGCTATAAA GCATCCAGAG
29501 AAGTGAGAAT TTGGTCCTGC TTTTAAGTAG CTTATAGTCT AATTAGGGGG
29551 AAGTAATCAG ATAGAAAGGA AACTAACAAT ATGCAAAAGG AAACTCATAG
29601 TTTGTGGTAA ATGCCAGGTG CTGCTGATAG TGGCTTCAGA GAGATCTCAT
29651 AGATGCTATA GGAGGTCAAA GGAGAAGCGT GCAGCTTGAG CTAAGTTTTC
29701 AGGGAAAAGG GTGAAAGAAT TAGTCATTAA TGTACACCTA CATTACCTGC
29751 CAGACTCCAT TCAAAAATAT TCTTACCAAA TCATCACAAT ACCTTGTTGG
29801 TAGGTACTAT TACTATTTTA CAGAGGAGGA AAGTGAGGCA AAGACACATT
29851 AAATAATTTT CCCAGAATCC CAAGGTGTGA GGTGGAGCAA GGACACAAAT
29901 CCATGGCTCT AAGTCCCTCC TAGTATATCC TGCAAACACA TCTGGAATTA
29951 ATGCAGAGAG GAAGGGGAGA GGCAGTGTTC TGCAGGAGTT CAGAGCCATG
30001 ATAACCCTTC TTGTGTGGCT TTTGGTAAGT TATTTTACCT CTTACCCTCT
30051 GTTTCCCCAT CTGTTCAATG AAGGTTGTAT ATACACACAT TATATGGCCG
30101 CTGTAAGTGT GCAGTGATAT GATGCATGGG GACTCAGTTC ATGAGGCAGT
30151 GTGAATTCTG AAGGTATCAC AATGGGACAG GTGTTTTTTT CTCCACTCAT
30201 TTTCTCCGAA AGTCTTTTGT TTTGTTGCCC TCCCTCTTTG GGGCATATGC
30251 TTTCAGCTCA TACCTTAATG ACATCAGAAT CTGCAATTTC CTGGCAACTT
30301 TTGTGGTTAA AATTATTCTG CCCTTCCATT TTAAAGCACT AATAGCAAAG
30351 GTATTAGGTG CAAAATGATG ATAAAAATAA TTGCAATTTT TACCATTAAA
30401 AGTCATGGCA AAACCACAAT TACTTTGGCA CCAGCTGAAT ATTTTGAAAC
30451 TCCCTACTCT GATGTTAACC AAGTTCATGA TTCAAAGAAC TTGCAGAGGG
30501 GTAGGGGAAT TTCAAGGGAA AGGGGGAGAT GCCTGGGGTT GTCACACACT
30551 CTGTCTTTCA TCCTCTATTG ACATGTTGGT TATTTGGAGA TGGTATTCAG
30601 TTCCACTATA GCCCCTCAGT CACTGTAGAC CCTCTCAAAG GGGCAATCAT
30651 GTTTCCCTTA GGTCAGGTCC ATTCATCTAA CCCCTCTCCC GGGGGCATCA
30701 CCTTGTTTGT TCCAGCAGCT GTCTGGCCAA ACTCACACCT CCTCCTCACC
30751 CTCTAGCCCT TATGATCTGC TTTGGGGAGC CATGGGAACC CCTAGTTTCC

30801 TCTTTCATAC CCACTGAGAT TCACAAGTAA CTAAGGTCAA GGCGGGGCTT
30851 CATTGCCTTT CTGCAGATAC CTTACGCTAC TGTTCCTCCT CGCCTGGCTG
30901 GCTCCACACT CCAGCAGACC TTCTGCTGGG CGAGAAGCTG CAGGCCTGAA
30951 TCTCTGTGTT CTCATATGGC CCCAACTCTT GGGATTACAC TAGCTCTTGT
31001 AAGAACTCAA TGCTCTGCTC TGCTCATTTT GATGCCATCA AAGAGGGCTT
31051 GCAAGTTACC AGCTGGGAGT GAACACCAGT GTCCTCTTTT TAGAGGTACC
31101 CCTAATCTTT CTGAACAATT TTGCTGGCAC CCCTTCACTT GGCTTTGCCG
31151 GGGTAAGAGG GGGCACTTCT CTCCTTTCCC TCATGAAAGG AGGGAGAGAA
31201 GCCAAAAATC TCCCTACTAG TCAACAACTC AGGCACCCCT CCTTCTCTCC
31251 TCTATTTTAT AGACTGGGAA GGGAGTGATG GTTGTTGGAG GTGGCAGAGC
31301 CAGTTCAGCT GCCTTTTGTG AAGTCCTGAA GGAGGTGTCT ATCCTCAACT
31351 GCTGGCTTCT GTCCTTAAGC CTGGGGAGAA TTAAGTCCTC TTTGCCTCAG
31401 TTTGGCACTC CAATTGCCAA CATTGGGACA GCAGGAAAAG TTCCATCCAA
31451 CATCCCATTA AATATGTAAT GTGTATTAGC ACAGCGCCTG GCACTGGGCA
31501 GGTATTTTCT AAGTGATAGC CAATGCGAAG CCTACTTTAT TATTTTCCTC
31551 TTTGCTTAAC CTACAAGGTG TCTAAGACCA TTTGTTTGTC CACACATAGT
31601 AAGATAAACA GCACTGAGAC TGTGGTCCTT TCTGCCCTGT GTCCTTATCC
31651 CACCTGGGAA TCTGGAAAGC CAAGCCTAGA CACACTCGTT CCACAAATGT
31701 TTACTGAAGC TTGTTCTATT CAAAGCACTG TACAGCTACA AAGACCATCT
31751 TTTCTGAACT CCAAACCAGG CCACATGGTT GGAATAACTT CAAGTATGGA
31801 GACCAAGAGA AAAGGTGGTT GTTGTCAGCA AAGCTCTGAG TCCACACCTT
31851 CCAGGAACTT ATAGTTGATG CAATGGTGGG AGAAGTCTGA ACCTGGATTC
31901 AATCTGCTTG ATTCCGATGA ATGGTGCAGT AGGCAGAGCC ATGAGTTCAG
31951 AGCAGGAAGA AACCACTGGT TCAAAGAAGC ATCTGTCACA TCGAAGCTGC
32001 TTTATAGTCT GTTGGGAAGC ATGCATAATA ATTTATTCTT TCTTTCTTTC
32051 CTTTGGTCAA CAAAGATTTC TTGAGTCCCT ACTATGTGCC AGGTACTCTT
32101 CTAGGTACTG AAGATGCAGC AGTGAACAAA GAAGATACAA TCCCTGCCCA
32151 GCGGAGCTTA CATTCTAGTT ATCGAAAGTC CCTTTCTCAG TGGCTGCTCT
32201 CTTTATTTGA GAAACCATGG GCTGTTCTCC TCCCATCCTA GGGCTGCTGG
32251 CTCCACAGAG GCACACAGTC CATCAGGATG CTCTGCCAGC CACCCACCCA
32301 CTCAAGACCA AGGGTTACGC TGTCAGTGTG AGCAGGGACA CTCCCGTCTC
32351 TGCTACCTCC TTTCTCCTGA AAACAAGATC TCAGGGAACA TCTGCCATCC
32401 ATTTTCCCTC CCTGGGGAGT GACAGGAAAG GTGTATGGAG GAGATTGAGC
32451 GGAGTGATGG ATTGAGGCAC TGTGAAAGTG AATCATTGCC TGACATGGGA
32501 ATGAGGAGAC TTGCTTAAAG GACAAGCCAT GCTAAGTCAT CCATCGTTCT
32551 CCCCTAAGGA GGTGAATTGA AGTTCCCATT TTTCCCAGGG AGCCAAATTA
32601 ACAAGGTGCT GGGAGATTTC CAAATTAGAA AAAAAAAAAA AAAAAGGCAC
32651 CACCAGCTCT CAAATCAGAG AGGCTGTTGA GTTGTTTTTT GGAGCAGATC
32701 ATTGTATTTG GCATCTAACC TTGAAATAGA GGAGAAAGCA TGGAATTTCT
32751 GCTGAAAACT CATCCTTCTC TGAGCAGGTG GTACAAATAA GCATCGTTGT
32801 GTTCTCAGAG GCAGGAACCA CATTTGCACC TTGATACCAA CTACCTCAAT
32851 AACCACAGTG CTGAATTTTC ACAAATTGCG AATTAGGAAA TTGTTGCTCA
32901 TTTTACAATT TGGTTTCCCT CAGGATTCCT TTTAAGTAGC CAGCTACCCC
32951 AGTACTTTTG AAATATGACT TGCTTATAAA AATTTGATAG GCTTGGCACG
33001 GTGGCTCACA CCTGTAATCC CAGCACTTTG GGAGGCCGAT GTGGGGTGGA
33051 TCACGAGGTC AGGAGTTCAA GACCAACATG GTGAAACCCT GTCCCTACTA
33101 AAAATACAAA AACTAGCCAG GCATGGTGGC ACATGCCTGT AATTCCAGCT
33151 GCTCGGGAGG CCAGGCAGCT AGGCAGGAGA ATCACTTGAA CCCAGGAGAT
33201 GGAGGTTGCA GTGAGCCAAG ATCATGCCAC TGCACTCCAT CCTGGGTGAC
33251 AGAGCAAGAC TTCATCTCAA AAAAAAAAAA AAGATATATA AACAAGTTTT
33301 TATAATATTC TCAATATGAA CTAGTAGAAA AAAAGCATGT GTTTTTAGGT
33351 CTTAGAGGCC TGGTTCCCAG TTTTATCTCT GACTCTAATG AGGTATAGTA
33401 TTACCTACAT TGATTAGCCC TTCTATACTT CATAGGAGAT GCTCCAAGAC
33451 TGCTAGCTTT CTTCATTCAA TAAAGAGAGA TATAACAGGA TGGGCCTTAA
33501 AAGTAGCATG CATTTCTTCT TTCATTCACT CATTCAAAAT ATTTTCATGC
33551 GTGAAAATGC CAAGGATGTT TGGTCAACCA ACTCTTCCCA GACCCTGGCT
33601 GTGAGCCTGG CTTAGAACAA TTCCATTTTA ATGGTCCATG CCCTCAGGCA
33651 CTTGTATTCT AGTAGAAGAG CAAGGTAAGA AAACAGCTTA AAAAGTTAAA
33701 CAGTTTTAGG TTGAGATGGG TGTTGTGAGA AAAATAAGCA GGATGCTTTG
33751 AACCTATGCA GGTAGGAAGG TCTGGAAAGG CCTCTCTGAT ATGGTGATGG
33801 TTAAAGCAAA ACCAAAAAGA CCAAGAACAC ATGGAACACA TGAAGGGCTG
33851 GAAGAACAGT GTTTTATGGG GAAGGACTAG TACACACAAA GGCTGCAAAG
33901 GCGAGTGGGC TCATTATGTT CTAGAACATG CCAAAAAGCG GGTGCAGCTG
33951 GAGAGGGAGT AAGATGGCAC AAAAGGTGAG TGAGGTGGAC AGGAGCCTTA
34001 TCACGCAGGC TTACACAGGC TCTCAGAAGC CCTGCGTGTT GGTTTCTTGG
34051 GACTACCGTA ACAAAGCTCC ACATACTGGG TGGCGTAAAA CAACAAAAAT
34101 GTATTGCCTC ACAGTTCTGG AGGCCAGAAT TCCAAAATCG GGTGCTGGCA
34151 GGGCTGCGCT CCCTCCAAAA CCTGTAGAGG AGAATCCTTC CTTGCCTGTC
34201 CCTAGCTTCC AGTGGGTTGC TAGCAATCCT GGGCTGGGTG ACTCCAGCTC
34251 TGCCTTGGTT GTCACAGGGC GTTGTCTTTG TGTGTCTCTG ACTTCACATA
34301 GCCCTCTTCT TCTTCTTTTT GTGTGTGTCT GTGTGTGTCC ACTCTGAGGC
34351 ACAGAAGTTT TTATTTATTT ATTTATTCAT TTATTTATTT CATTGATAAA
34401 CATAATAGTT ATGCATAGTT TTGGGGTACA TGAGATATTG GATACATGTG
34451 TACAGTGTGT GATAATCAAA TCAGGGTGAT TGGAATATCC ATTCACCTCC
34501 AAACATTTTC TCATTTCTTT GATTGGGGAC ATTATAATTC TTCTAGCTAT
34551 TTTGAAATAT ACAATAGATT ATTGTTTACT ATAATTTCCC TGCTGTACTA
34601 TCGAATACTA GAACTTATTC CTTCTGTTGA GGGTGTACTT TTGCACCCAT

FIGURE 3, page 9 of 57

34651 TAACCAACTT TTCTTTATGT CCTCCTTCCC ACTTCCCTTA CCAGCCTCTG
34701 GTAACCACCA ATCTACTCTC TACCACCATG AAATCAACTT TTTTTTTTTA
34751 TAGCTCTCAT ATATGAGTGA GACTATGCAG TGTTTGTCTT CTGTGCCTGG
34801 CTTATTTCAC TCAACATAAT GACCTCCAGT TCTGTCCATG CTGCTGCAAA
34851 TGACAGGATC TTATTTATTT TTTTATGGCT AAATGGTATT CCATTTTGTA
34901 TGTATATCAT ATCTTCTTTA TCCATTCATC CACTGATGCA TATTTAGGTT
34951 GATTCCATAT CTTGGCTATT GTGAATAGTG CTCCAATAAC CATGGAAGTG
35001 AAAATATCTC TTCAACATAC TGATTTCCTT TCTTTTGGAT ATATACCCAG
35051 TGGTAGGATT GCTAGATCAT ATGGCAGTTC TAACTTTAGA TTTTAAAGGA
35101 ACCTCCATAC TTTTTTTCCA TGGTGGCTGT ATTACTTACA TTCCCACCAA
35151 CAGCATATGG TCATCTCCTT TCTCCACATC CTTGCCAGAA TTTGTTATAT
35201 TTTGTCTTTT TGATAATAGC CATTCTGACT GGGGTAAGAT GATATATCAC
35251 TGTAGTTTTG ATTTGCATTT CCCTTATAAT TAGTGATGTT GAGCATTTTT
35301 TTATATACCT GTTGGCCATT TATATGTCTT CTTTTGAGAA ATGTCTATTC
35351 AGGTCTTCTG CCCATTTTTA AGTGGATTAT TTGTTTTTTT GCTACTGAGT
35401 TCTTCGAGTT TCTTATATAT TCTGATACAC AGCCATCTTC TTATGAGGAC
35451 TCCAGTTATA TACGATTAGA GAGGTCCACC CTTTTTCAGA ATGAAATTAT
35501 AGCTTAACTA ATTACATCTG TAGTAACTCT ATTTCCAAGT AAGGTCACAT
35551 TCTGAGGTAC AAGGGTTTAG GACTTCAACA TATGAATTCC AGTGGGACAC
35601 AGCTCAACAC ATGACACCAT GGTAGGGAAC TTTATTCTAC TTGCAAGTTC
35651 TGAGTGTCTT ACGCAGGTAG ATGGACTGGT GTGATGTATG CTTTAAAGAC
35701 CGCTGTGTGA AGATGGCCTT AGGGTGATGA GGATGGAAGT TGGAGACTAA
35751 TAAAGGACTA AGAAAATGCT AAGAAAATCC AGGTGAGAGG TGATGATGGC
35801 AGAACTAAGG TGATAGCAGT AGAGAGAAGA GAAGTGGATG GAGATTAGAC
35851 ATCTTTTGCA GAACGAATGA CAAAATACCC CTATGGATTG GACATGGGAT
35901 GAGGAAAAGG AAGGACTTGA GGGTGGTGTC TAGGCTTTTT ACTTTAATCG
35951 TGAAGGGAAG CTGGTGCCAT TTACCTTGTT CGGACAAACC TGGAGAGGAT
36001 CAGGTTAGGG ACTGCGAGTG GTATGGACGG CAAAGGAATG GGAAGAATGC
36051 AGGGATTAAA AATTGGAAAT CCCCCTCCCC AGTCAACAAT ATCTTACTTT
36101 TATCTGAAAA ATACTAAGTA AAAAAGCATC CTTTTGTTGG AAAGCTCAAT
36151 CCTTGTTAAA ATGAAGACAT CTCTGGGAGA GGAAACATAG TGAGCACCTT
36201 TCCCAAAAGC AGCCACTGAT TTGGAGATGA GACAGAGTAG CATACAGGAC
36251 ATCAGAGAGA ACATGCTCAG GACAGAAAGA GCAATGTAGG ACAAGGCAGT
36301 GTCTTGGCAT CACAGTCTTT CCTCCGACTG GCTGTGAGCA AGTGCTCAAT
36351 TTAATTCCAT CTCAGTGCTG GGTCAGGACA AGTGCCCAAA AGCAAATTGA
36401 CAAAAGTACC AGCATGATGG AGTTAGAAGG TAGCAAGTTC CCTCCACAGA
36451 GCCCAGCTGG AAAGGAAGAT AGAGGGGAAG TTGACCCCTG GGGATGGGGA
36501 ATAGGGTGAG AGGAGAACAT GAAACTGAGA AAAGGGCTTT GAGTGAAATC
36551 TAGGCTAAAA GCTAAGGTTT CTTTAGAAAC CCACCATTGA CCCAACATGA
36601 CCAGGGCTTT CTCTTGACTT GATTATTTTT GATACCCCAT CTTCTTCTGT
36651 ATTCCTGGAA CTAGCTCTCC CAAGCCCCAG AATTGTGCTT CTATCAGAGC
36701 TGGGTTTTCA TCAGAGTCTC CCCTTTATCC TGTATCTCTG TTGCCCTATT
36751 TTGTTTGAAT TCCTGCCAGG TCAGCTGAAT TTGGGCATTT GGGGTGAAAA
36801 ACCATCAAGT GTGGCATCCT GGCTTTGGCA CCTGGCACAG TGTGACCCCA
36851 CTGGTCTCTC CCTCACATTT GCTGTGGTCC GTGCACGGAA TTTGTCAAAA
36901 GACCTCCTCA GTATCAGCTT TCCTGCAGCC TCAATGCACC TTGTTCTGAA
36951 TAGGATATTA CCCCCCAAGA GTATATTAGG GCATTTTCCT ATGCCAGAAG
37001 GGGTCCTTAG GCCTCTTGCA GTTTTTCTG GGTGACAGTG AAGGAGGAGG
37051 TGGCTGCAGA GCTTACTGCC TGTGGACTGA CCACCCCAGG GCCTGGTGTC
37101 AGGACCATTT GTCCAGCCTG TTGAGTGAAG GTCATTCTGC CTAAACTGTA
37151 AGCACAAGAG AGAGTTCAGC ATCATTTGCA TCCTATTTTA TTGTCTTTCT
37201 TCTCTTTTCT TTCAAGGCCT CATTTTTTTT GGCTTGAACA AATGGTAAAG
37251 GCCATTTTAT TACAGGTACC AAGCCAAACT TTCCTTGGTT TTGTGGCCAT
37301 CCTGCTGGGG AAGGAAGTAC TCCTTTACTT TAAATAACTT TAAAAACATC
37351 TGTTTGGTCT CAGGGGCTGC AGCTGGAAAG ATTTTCTAAC TAATACTTGT
37401 TTTATGGGGG TGTTTTTGGG GGGGTTTATT GAGTGTCAAA CCTGGCAGTA
37451 AATTAGAATC AGAAGACAAC AGTTAGTGAT AAGCAGAGAA GCCAAGGATG
37501 TTACCATAGG CAGGCAGCAG AGAGAGGGGA ATTGGTGGCT GGCCCCCCAA
37551 AAACAGATTT GAAGATCTCC TTCTGTCATG TAGTGAATCC CCAAGTGCCT
37601 AGGGTGGGCT GTGATTACTT GAGCTCCTGT CTCCACTGTC TCAGCTCACT
37651 TGCCTTGGGG TGGACACACA ACACACATTT GCTCATAGCA TCAGGTATTC
37701 AGGAGCAAAG AGCTGAATTT ATCTGGTTAA TTTAGATACC CCTACCCCCT
37751 CTTTTAACAC CAGATTGCCA GGATCATGAC CTCAAAAGGC TACCCTGAAA
37801 TGCAATTGAC AAATGGGATG AAAGATTTCC CGTTTCATCC ACATTTGCCT
37851 CCTGAGCTAC TTACAGCAGC AGGTCACCGC AGCCAGAGCC CACCTGCTTG
37901 CCCACCATGC CCGCACACAG ACAATGCTGC TTCTGTGGCT GGAGGTCGGA
37951 ACACCTCAGC ACTATCTCAG TTTGGCTGCA GATCCTCTGT GTGCTTGGTA
38001 AACAGGTTTC CTCATCTGTA AAATGAATTG GCTCTTCCAC AACTTTTTTA
38051 AAAGCACTAA CATATTAGGA CTCTCACTAA ATACTCAAAT GCTAAACTCA
38101 AATACTAAAA GAGTGCAAAG GGATGGGCTC CCAAATATTA CAGTGAAGGC
38151 TGCAGCATTT TCTGACCTTG CTGCTTTTTC TGGTGAGTGG CTTTTATTTC
38201 TTAGTTTGGT TTCTTCTCTC CCATTCTAAT CAAGCAAGAA GTGACCACCA
38251 AAAGGGGCAC TCACCAAACC AGAACAAGCT AGTTCTTTCA TCTTTAATTC
38301 ATTGCAACCA AACAGATGCC ACAGAAAGAG CCAAGGGCTC CAGGCTTTAG
38351 CTCCAGCCTT GCCATTAACT ACATATGTAA GTCAGCCATG CTGGTCTGCA
38401 GGTTCTTGCT TTGCATGATC AAGGGACAAC TTGGAAGGTC TCCAATCACT
38451 CTATTCCCCC AGATGGAAAT GTATTCACTT ATTTCCTGGA GATGTCTGTC

38501 CTCCTCCCAG TTAAAGACAG ACCTTGACCC ACCTCCACTT CCTTCTCTGT
38551 GGCCCTGTCT TATCTGTCCT CTTGTTCTTG CCTCTTCAAT TGTTCTCTCA
38601 CCGTGTTTGT CACTTCTGAG CTATCACTGT GATCCCCCTG ATTGTTTTTC
38651 TAATGTCCCT GAACTTCAAC CTGATTTTCA CGCATATACA ATGTCTTCCT
38701 AAACACTTAT AGACTCTGAC ACATTCTGTA ACTGACACAT TTCCCTTTAT
38751 CAAATGCAAT CTAAGAAGCT CACAGTTTCT CTCAGTTTCA ACAAGAGAAA
38801 TCAGGAGCAC TTGAATTATA CAACTTGACA TTATTAGGGC TGATGTCTGA
38851 TTTTGTCCTG TCTGCCCCTG TCATTTCTGT ACTACCTTTT ACAAAACCTC
38901 TCCTATGACC TGTGTCCTCC TCCAGCTCCA TTTGAGAACA CCTGCTGTAT
38951 ACCCTGTGGG CTAGCTTTTA TTATGTTCGC CTCAATGATG AAGAAACAGG
39001 CTTGGAAGTT AAATTATCTA CCCCAGGCCC ACAGCCTGGA ACCTAGGATT
39051 CCAACCAAAC CTTGTCTGAT TCTAAAGCAT AGCAGAGGCT CCATACTCTG
39101 CCTCCCTCTT CTACATCATT TCAGTTTCTT CACTTTCCCA CCTCCAATTC
39151 TCACCCAAAC TGAATGTCTC ACAGTCTCTG TGCCCCCACT TTGCTCCATC
39201 CCTTGGCCTT CTGCAGTCCA AGCTCCATTC TGAGATCATC CAAGGCTTCT
39251 CTTCTGTGTT GATCCTTGGC CTTCTTGGAG TCTCTTTCTC CCATGTTCTC
39301 CACAACAGAG CATTCTCCTG ACTGTTTTCA TTCTGCATCT CACTCTTTCA
39351 TCAGTATCTT TTTCTCTACC ATGCCCCATA AATTTGGGTG CTCCTGAGGG
39401 TCCTGTCCTT GTCCCCTGCT TTCTTGTTGT ACAACCTCCT TGATCTACTT
39451 CATCTACTCA AGTTTGGTCC ACAATTTCTA TATTGTGAAG ATTCAAATCT
39501 GCATCTCTAG CCATATATCC ATTTGCCTGC TAGGCATTTC TACCTGAATA
39551 TTTTATAGGC ATGCCAGTGG CTCTTACTCT ATGGCTCTTA CTCTAAGTCT
39601 AGACTACAGC AGAAAGCAAT GCTCTTTTTA TTAAGGCATA GTGCCTCTTT
39651 CAGAATAATT TACAGCATAC AACCAGGCCT GCTGTGCAGC ATTACAATTT
39701 GTCATTAAAA CTCCATTCCT CTTGCCAGAG TAAATGAGCC ATTTACAGCC
39751 AGGGCGCCAA GATGGACTGT TGTTATTTTT TCTGCCTTTG TATTATGAGT
39801 ATTCATGGCT CTCCTCAGAC AAGCTCCTGG GGATTCCCAG TGGAGTTGCC
39851 TTAACATGCA GGTCAATTAG CCAGGCTCAA GGGTAGTTTC CTGGATATTG
39901 GTATCCCCCT TGCAGAGGAC TGCAGGAAAG CTGAACAGTG TTCCCCCAAT
39951 GTGGGTGGTG ATCCTGAGAA ATATCATTTG TATCTGCATG TGCTGTCTCA
40001 CACACACTAG CTCACATGTG CACACACACG TGCATGCACA GGACAAAACC
40051 AAACACAGGG CAACCCAGCA TCTGCCCCCC AGCCATCAGC ATTGTTACAC
40101 CTTTATAGGG GGCGGGAACA GGTTGGTCAG CAGGTGAACG TCAGGTGAGT
40151 TGAGAAAAGT TATTAAATCT TAAATCCTTA AGGAAAGTTA TTAAATCTCT
40201 TCTAAAATGC ATGCATAGGC GGGCTCAGTA ACTAACATGC AAATGTTTAG
40251 GGTCTGAAGC TCCTACCGAT AATCTTTCAG ATCTCAGAAT TCCAGCCCCT
40301 TGTGCTGTTC TGGGTTGTCT GACACAGACG AAGCAGAGAA CAGTAGAATA
40351 AACAGCTCAG TAAACAATTC ATTGAGGGAA AGAGAGTGAG AAGATTCACT
40401 GGACAGCTAG AGGAGGAAAT ACTGCTGGTG ACTATGGAAG AAATTTGCCC
40451 TAAGGCCTGC AGGCAATAGC TTGGTCTTAT TTATCCTGGT GTCCCACCCT
40501 CTCCTCCAAC ACATACTGCC CTGGCAGGTA CGTAGAAGAT GCGTGAAAAT
40551 ATCTTTTGAA TTGAGCTATG CAAAAAATAC TGGATTCTGC CCTCCAAGAG
40601 TTTACTGTTT AGTTTCACAG AAAGCACATG CCCTCCTTTC TCTGCCTCTT
40651 GAAGACTGAC CTATCTTTCA AGGCCACTGG CCCAATTCTG TTTTCTAAGT
40701 AAGACCACTG AGTCAGTGGT GACCTCTCCT TCTCCCTAAC AAAGTCTGAT
40751 TTACTTGAAT ATACAACTAT CTCCCTCTTG GCCTGTGAAT TTCTTGTGTT
40801 AGGGAACATA TCTGATTTAT CCTTATCTCT TCCACAGTAC CTGGTGTAAA
40851 ATGCCCAATA AATGCATTGA AATATTCATG AAGCTTACTA AATGCTCTGC
40901 CTTATGAGCC ATGAAATATA AAGTGCCTTA AACTTTGTTT TTCTCTTATG
40951 TAAAATAAGG ATAATAATAA TGACACCCCT ATAGGATTGC TGCAAGGATT
41001 AAGTGTGATA ATATATATAA AACTCTTAGC ACAAACACCT GGCTCACAGG
41051 AATAGTAGCT ACTACCATAA TGGTAACTTC GAGGGCAAGT TTTCTCAGAG
41101 TTATTTAGCC CTCCTTCACC CTGTGTCCAG GAGTGCAGAT CAGAATGGTC
41151 AGATTCCAGG ACACCAAGTT TTCTGTGGGA GCTTCCCTAG GAATATAACT
41201 AAGGAATTTA AATCAGGTTC AGCTCATGCT GTTACACTCT CTTCCTCCAC
41251 TCAGGCATTG GGTGTGGCTT TTCCAAGCTT GAGAAGGGTG TGATCTGAGA
41301 TGGGCTTGGG TATAGAGGGG AATTATATTT AGGTCTACCC TGTATAGGAA
41351 AAAGTGCCTT CCCAAAGTCT CCCTGGCCTA AAGTATAAGA GATATGTGTT
41401 GGGATTTAGA CCCAGAGCCC AAGCCAATAA TGGGACCCCC TTCTCACATG
41451 TGGCTACCTC CTGCTATCAC CACAACAGCT ATCATACCCA TAACTACAAC
41501 AGAGGCCAAT TAACGTGGTG ATAATTGACA AATGTCAAGA CATCCTACAT
41551 TGAGGCACAC TGTGCGTTTT GCGTGAGCTT TTAAATTGGT AGGGAAGGAA
41601 AACTTTTATA CCTACACCTA TCATGGAAGG CAGAAGGTAA GAGCTAAAAT
41651 AAAGGTATGC CAAGAACAAA GGCAGGAAAG AAGGGTTTTA ACAACTTGAG
41701 GCCTGATCCA TTGATTAGTG AAGAGGAAAC ATGTTCAAAA ACCACTCTAT
41751 AACCACCTTC TCCAAGTTTT TTATAATTTT GCTTCTTCGG ATATCTTCTC
41801 ATCATAGTCT TAAATGCCAT CAAATTAACT GAAAAATGCT AAAAATGCAA
41851 CCACTCTAAG AGAATGGGTT AGATGGGAGA TGGCTTTGTT AAAGAAGTCG
41901 GTCTTAAAGC AAAAGTAGGG CTTTGTCATG GTAGTATGGA AGGAAGGACA
41951 TTTTTGGTCA AGAGAAGAAA GTGCAGGGCC TGTTGAGGAA GGAATGAGTA
42001 GTAAAATATG GCTAGAACAG GGTGCAGAGG GGAAGAACTT CAGAGAATGA
42051 CCAAATAAAC AGGCTGAAAG GTGTAGACAT TATAGGCAAT AAAGCAACCA
42101 CAGAGGTTTC TAAGCCATAG GGTGACATGA TAGATCTGTA TTCTAGAAAA
42151 GTTAGTTTTG CAGCAGTTGT GTCCATTGAA AGGGACAGGA TAAGGGAGAT
42201 AGATAAGAAG ACATGCTATG ATGATAACTA GATTTGGATA CCAAGTGGTA
42251 TGGTGGAAAG GAATGAGAGA ACAGGGTCAC AGATGAATGA CTGCCCAATT
42301 TCAATCCATC ATAACAGGAT GTATAGGATT GCCCTTAAGT AAGATGGGGA

FIGURE 3, page 11 of 57

42351 ATCCAAAAAC GAGGAACAAG TTTGTAAGGT TTTGGGGGCC AATGATGAAT
42401 TCCATTTGGG ACATGTTGCT TTGGATATAC CAATGGGACA TTCATGTGAA
42451 AATGATCTCG GCAATCCTAT CCTGGAATTC AGGATAGGAT CAGAATGAGG
42501 GACACAGTTT ATAAGGTAAA CAGAATGGAG GTGATATAGA AGATAAGGGC
42551 ATAGATGAGC TTACCAAAGG GGAGAGTTTA GAATGAAAAG AAAAGACCAA
42601 AGGCTAAGCC TGTGCTATTC TTTCTCCTCA CAATACGCTT CAGACCTGGG
42651 CACAAACCAT CAGTGAGTGT CATGATAACA CTACTGTGGG CAAATCCCCC
42701 CTCTATAAGG GCCTGATTTC CTCCTCTATA AAATAGAGGG TTGAACAGGG
42751 TGGTCCATAT CCTGTTAATT GTGTTTGGAG AGCACACAAC AAACCAGCTA
42801 CTATCCAAAG GGGACATCCC GAGGCAGGAC TAAGCAAAGG AAATCCAGCA
42851 CAGGGAAAAC ACTTTCTGGT GCTGGTCCCA GTTAGGCAGC GTTCAGTTTA
42901 ACCCATCACC ATCACCATCA GTAGCTTTCA GCTGCTACTG ACCACACTTA
42951 TAGGAAGAAA AACAATTAGA ATGGAGAGCT AACTCTTTGG AAATGGTCAA
43001 AGAACACGGG TCTACAAAAC CGTCAATAAA GCGCTAAGAT GCCTGGGCGG
43051 GGTCAAAAAG TCTACCTGGG CGGGGTCAAA AAGTCTACCT GCTCAGCATA
43101 TGGGGCCCAG ACATCTGACC TTTACCAACT CCACAATAAC CACTTCATCT
43151 ATGGATCCAG TCTTGGTATC ACCTAGTCGC TGTTTTCAAG TAACAGAATA
43201 TTTGGTTCTC AATGGTAGGT GACTGGAATA CAGCTTACTT TCTCCCACCC
43251 CTACCGCCAA TCCTTTCTGC CCCCTTATAG TTTAATTTGC TTGTAAATTA
43301 CTTGGGAATA CATTTGGGAG CCATTATAGG GAAATAGAAG GCAGACATGA
43351 TGAACAGAAT GCAGGGTGTT TTTTATTACT TCACATTGTG CTCAACAATT
43401 AGGAGGAATT CTAGAAGCCC CTCCCAGTGG CCAGGAATTG GTCATAGCAT
43451 GAATAAACTC AATATAGGTT GAGTATTCCT TACCCAAAAT GCTTGATACC
43501 AGAAGTGTTT TTGGATTTTG GATTTTTTTT TTGAATATTT GCATTATATA
43551 CTTACCAGTT CAGCATCCCT AATCCAAAAC TGAAATCTAA ACTGCTCCAA
43601 TGAACATTTC CTTTGAGTGT CATATTGGCA CTCAAAAGGT TCCAATTTTG
43651 GAGCATTTTC AATTTTGGGT TTTGGGATTA GGGATACTCA ACCAGTGGTA
43701 GGTTTGGGAT GATATCAGCA TGCTAAGGTC AAAGAGACCT AGCTGGGAAG
43751 GGTGGGAGGA ACATGGAATT TTCATTCTCT GGGCACCCCT TGAACAGTCT
43801 TACTATTAGG GCCCCAAATT TGTTCTAAGT GTGTGTGTGT GTGTGTGTGT
43851 GTGTGTGAGA GAGAGAGAGA GAGAGAGAGA GAATTTTCTT TCTTCCTTTA
43901 TATTCTAAGT TCCTCAGGAC AAAATTTTGG GTTTCTTTGT ATTCTCCCTG
43951 CAGCTCCTCA TGTAGTTCTA AGCAAATAAA GGAATTCATT AGGTCCTTGA
44001 TTTCAGAAGC CTCCCAGTTC TCTATGTAGG AGGAATCTTA GGGTGGCAAG
44051 ATAAGTTGAG GGACTTTTCT TCAAGCACAT TTCACAAGTA AGAGAAAATG
44101 TTGACTGTGT ATATCTAAGA ATGGGTGGGG CTCAATGATG CCCCCCTAAG
44151 TTACTCTTTA CTATTATTGA TTGATTGATT GATTGATTGA AGAAGCAATG
44201 TTTTGATTGA TTGAAGAAGT AATGTTTCCA ATGGCTACAG CAGACTGGAG
44251 CAAAAGAACA AAATGAAAGA AAATACATTA GGCTTTCCAT TTCTTCTAAT
44301 TCTGGGGCAT CTGATGAAGC TTTGGATCCC CCAAGGTAAG AGCTGGACTC
44351 TGCTGGTGAA AACTCTTTAG GAAAAACAAA AGAATATTGT CAGAATCTGA
44401 TGCACCTTAG AAATGATGCA GCAGAACTGC TTTATTTTCT AAAAGGTGAA
44451 ATGGAGACCC AGAGAAGCAA AGTGATTTGT TCATGATCAT ACAGCTATTC
44501 AGTAAAGCCA GGACTTCTGT GATCCACTGT CCTTTCCTTA AACCAGTGGT
44551 TCTCAACCTT GGGAGCTTTA AAAAACTGCT AGTGTTGGAT CCATCTCAGA
44601 CTAATTAAAT CAGAACCCAT GGGGATGAGG CCCAGACATG AGTGGGTTTT
44651 TTGTTCTTTT TTAAAAAAAA GCTCCCTAGG AGATTTCTCA AAGAACTGAA
44701 AATAGAACTA CCATATGATC CAGCAATCCC ACTTTTGGGT ATCTACCCAA
44751 AGGAAGATAA ATTATTATAT AAAAAAGATA CCTGCACTCA AATATTTATT
44801 GCAACACTAT CCACAGTAGC AAAAATATGG AATCAACCTA ACTGTCCATC
44851 CATGGATGAC TGGATAAAGA AAATGTGTAT ATATACACAC ACAATGGAAT
44901 ACTATTCATT CGTAAAAAAG AACAAAGTCT GTCTTTTGCA GCAATATGGA
44951 AGGAACTGGA AGCCATTCTC TTAAGTGAAG CAACTCAGAA ACAGAAAGGC
45001 AAATTCCACA TGTTCTCACT TACAATTGGG AGCTAAATAA TGCATATGCA
45051 TGGGCACAGA GTGTGGAATA ATAGACATTG GAGACTCGGA AGGGTGGGGG
45101 GAATGGGAGA GGGTCAATGA TGAAAAATTA CTTAATGAGT ACAACGTACA
45151 TTATTTGGGT GATGAATACA CTAAAAGCCC ACACTTTACC ACTATGCAAT
45201 ATGGCCATGT AACAAAATTG CCCTTACACC CCTTAAATTT ATACAAATAA
45251 AAATAAATAA ATAAAAGCTC CTTAGGGCTG AGAACTACTG CTCCTGTCCT
45301 ATGGGTCCCC AGCTTTATTT TAACTCAAAA TGAGTTTAGA AAAATTTATG
45351 AACCCATTTA AAAATATTTA TTGAGTATCT CCTGTGTGCA AGGCACTGTG
45401 TTATGTTAAG TGGCTGAAGG GAAATTAGAC TGGGGAAAAA GACAAGGTCA
45451 TGGCCTAGGT TTCAAACTAA TATAAAAGAC ATAACAAATA AGAAAGGATG
45501 CCACCTTCTT CCAACCCTCA TCCCTCTTCC TTTTGACAGT TGCAGATGTT
45551 GCTAATTCAT TTTGGCACCC TTTTTCTCTG ACCCAAATAT AGTCTTATAA
45601 ACCTTTTCAA CCCACGGCTC TAGGCAAGTA TCACCTTTTG CTCTTTTGGC
45651 ACCAGATCTC TTGAACACTA TTTACTGGTT TTGGAAAGAT TATACATGTA
45701 TGTCTGGAGT TGAATGACTG AACAGAGCAA TAATAAGAGT TAAAGCAAGA
45751 AAGACAGGCC TACAGGAGAT GGCAGAGGGT CTTGCCTGTC AGGCATTGAT
45801 TTTGAACTTC ATTGCATAGG CAATCAAGAA CTATTGAAGT TTTTGCACAA
45851 AAGACTATAG ATGAGATTAA CCTGGTTACC GTAAAGGACA AAGTGATTGC
45901 AGGTAGAATG AGGCCAGCTT CATAAATGAA TCATCAGGAT ATGAGAAGCA
45951 AGGGCTTGAA CATGAGAGGC CATAGTGGGA ATGGAGGGAA AGGGACAATG
46001 TGAGAAGCAG TGAAGGAGAA GGGCTGATTG AGTAAAGCAG TGGAGAAGAC
46051 AGTGAAAGAT GTCAGATGAC TACCATGTTT GGCGACTGAG TGAGGGAAGA
46101 GGTGGTGATG ATATTACTGA AGAGAGAGGC AAGGGGTGGT CACTGGATTT
46151 AGAGCAGACA TTATCAACTT GTGGTGTCCA GACATTTCAC CCTGGGAGAA

46201 ACCTGTTCTG AAGTGGCTTC AGCATCTCTG AGGTCAGATT CCTAGTTCTA
46251 CTATTTTTCT ACTGACTGAA ATGGAAATCG AGTAGGCAAG GCTTTTGATT
46301 TGTCTCAGTG GTCTCTTCTG TAAAATGGGG GTGTTTATAT CCATAGTCTT
46351 ATCACAGGGC TATTTGGGGG ATTAAGTAAG ACAAGTGTGG CAGAGCTTTG
46401 TAAACTGTAA TACACTGTGT ACAATTGGAT AATTATGGAT TCTTCTGACT
46451 CATCCACATG GATGTCTGCT GACCCTGGGG GACCGGAGCC TGGGAGGGAG
46501 GCCAGACCTG GAAATGGAAA CTTGAAAATG TTCTCTGTAG AAAAGATAAT
46551 TAACATTTGA GGATGGTTAA GTCCTCTTAA ATAGATGTCA GAAAAAATGG
46601 AGGTCATGTA GACAGAATGT TGGATAACAC TACTTTGTAA AATATTTTAT
46651 CTTATTTCCA TTATAAAAGA AAAAAAGCTG GGCTGGGCAC GGTGGCTCAC
46701 GCCTGTAATC CCAGCACTTT GGGAGACTGA GGCGGGTGGA TTACCTGAGG
46751 TCGGGAGTTC AAGACCAGCC TGGCCAACGT GGTGAAACCC TGTCTCTACT
46801 GAAAATAGAA AAATTAGCCG GGTGTGGTGA CAGGTGCCTG TAATCCTAGC
46851 TACTCGGGAG GCTGAGGCCG GAGAATTGCT TGAACCCAGG AGGTGGAGGT
46901 TGCAGTGAGC CAAGATTGCA CCATTGCACT CCAGCCTGGG CGACAAGAGT
46951 GAAACTCCAT CTCAGAAAAA AAAAAAAAAT AGACAGGAAA ATAAAAAAAG
47001 CCACCTCACA TAGTCTACTA CCACCAAACA CATCATTAAC ATTATATTTC
47051 TTTATTCCAT GCTCTTTGTT TTTAATATAA ACAATTACTT TTAAGGGAAA
47101 ATGAGAAAAG GAGAGAGTGA TAAGACTTTA TTTTAAAAGG TGGAATAATT
47151 CTAACCATGG AGAGTATTTA TAAATTTTTT TTTTTTGAGA CAGAGTCTCG
47201 CTCTGTCACC CAGGGTGGAG TGCAATGGCG TGATCTCAGC TCACTGCAAC
47251 CTCCACCTCC CGGGTTCAAG CAATTCTCCT GCCTCAGCCT CCTGAGTAGC
47301 TGGGATTACA GGCAACCGCC ACCATGCCCT GCCAATTTTT TTTTTTTTTT
47351 TTTTTTTGGA GATGGAGTCT TGCTGTGTCG CCCCAGGCTG GAGTGCAGTG
47401 GCATGATCTT GGCTCACTGC AAGCTCCGCC TCCTGGGTTC ACGCCATTCT
47451 CCTGCCTCAG CCTCCCAAGT AGCTGGGACT ACAGGCGCCC GCCACCGCAC
47501 CCAGCTAATT TTTGTATTTT TAGTAGAGAC AGGGTTTCAT TATGTTGGCC
47551 AGGCTGGTCT TCAACTCCTG GCCTCAAGCA ATCCTCCTGC CTCAGCCTCC
47601 CAAAGTGCTG GAATTACAGG TGTGAGCCAC CGTGCCAGGC CCATAAAATA
47651 TTTTTTATAG ACAAGTGAGA GCAGAAATCA CAGGTTCTTA TGAGCAGGAA
47701 AATTTGAAG GTCATCTACT CTGAACGTTT TTTTGTTTGT TTGTTTGTTG
47751 TTGTTGTTTG TTTGTTTTTG CTTAGTTTAC ATTTATTAAA TACCCGTTAT
47801 GGTCCAGGCC CTTGGCTAAG CGCCATCCAT GCAATATATC ACAAGATATG
47851 CCCAGCAATC CTAGGAGGTA GGGTTTATTA CTACCCATCG TACAGAGGAG
47901 GAAACTGAGT CATAGAGTTT TAGTGTCCTG ATCCTGGTCA CAGAGCCAGG
47951 AAGTGGCAGA GCAGGCCAGG CCAAGTCTGT CTGACATCAG AGCTCATCAG
48001 AGCCCTCCCC ATTGTCCTTG AACCAGTAAA GATGGAGTTC TTCTACAGGG
48051 GTGGTTGGGG GACAAGGACC CCATGGGTGT GTCTGAGTCA GAAACATCTG
48101 CGAGTGGGCT GAGAAATGAG TCTTCTGTGA AAAAGAGCAA AAGAAAAAAT
48151 GGGTCAGGAG CCAATAATCA TTGTCCATCT TTGTGTGAAT GTATGGTGTG
48201 GGAGTGGGAG CAATAAACGA TTCTAAGGTC ACACAGAAAA GATGCCACCT
48251 TCTCCAATCA CATACCGCCC CTCGTCCCCC AGTTTTCTCT GAAATAGCTC
48301 TTCTTTTGGC TCTATCCTGG CTTCTTCACA CAGGGGTGTC CAGTCATCTC
48351 ATCCTGGTGG GACAGGGATA GAGCTGTGGC AGTGGAGATG AGGAAGCTCG
48401 CCTCCTAAGT GAGTCTGAAT TCTTAAATAT GGAGCCACTC CATAATCATT
48451 TGGAGTGAAT ATTGGGCCAT GGCCCTTTTT CTTGCCAGCT GAGCTATGAA
48501 AAAAGGATGT CCTAAGACCA GAGGCTGTGG GACCATTCCC AGCCCCTGCA
48551 GGAATCAAAG GAGCTGACAG AATTGTTTGT TTGTTTTTTT CACAAATTGA
48601 AAAAAAAAAT GTAAAATTTT TGAAAAGAAA GCCTCATTGA AAAGAAATCC
48651 CTCTCCCCAG CTGGGCTCCC AGGCAGCCTC CTGCAGAACA TCCTTAGCAT
48701 TGCAGAGTTG TTCCCATGGC AACCGAGTAA GGGGCTTTTT GTTTTCCTTA
48751 GAAGATTGAA TCCTTTCAAC CAGAAGGTAA CCACTGGTTC TTCCCCACAA
48801 TCCACACTCC AAACCCCCTA CCCTTATTTG ACTACATGAC TAGTTTTGCA
48851 TTTATGGATT TTTTTATGCC TAATTGAAAA AGGCTAAATA TACAGAAACT
48901 GAGGCTGAAG TGGTTTAAGG AGGCAACTGG CCCAGTGGTT TCTCAGCAAC
48951 CACATGTCAA AGCTGTGGAC GTTAGACTTG ACGAGAGCAA GACATATCAG
49001 AATCTGTAGC AGGAGCATCT AGTCTCCCAG TTCAATAGTG TCCACAAAAG
49051 AAATCCAGAG GTTTTTGAAG CAAGGAATTT GGGTGGCACT GCTGTGAGAA
49101 ACAATCACCT GGCTCCTCCA TGGGGCATAG AGTGGAGATG CTTCTTCAAA
49151 TACCCCTTCC TTTCCAAGGC CATGACTCAG AATGACTGGC GTAGGGAGCC
49201 TGGACCTGAT CTCTTCAAGG AAGGGGAATC AGATGAGCTG TTTAATCTCT
49251 CTTGTAAAAT GAGGGGTTAT GAGACCATAG GCTCATTTTG GGGGGGGTCT
49301 AAAATGCAGT ATTTTTTGAA CTGATATGGG GAAAAAAAGA CATTTCTGAA
49351 TTGTTGTCAT GTTGCAGATT CTGGGCCGTT CCAGCATAAG CACCTTTCTT
49401 AGAGTACTTG GCTTTGTGAA GTAGTCCTTA TCCCCTCCTT CCACTATTTT
49451 ACATCAAGTT AAAATAGAGG AAGATGCCTA GAAATGGCCG TATAGACAGA
49501 GAAAACTGCA CTAAAACTCC CTCCGTCATG CCTGACTCCT CTCTAGACTA
49551 TGACCATCGA GGGGCCAGAA ATCATATCTT AAAGATCACT GTGCCTCCAG
49601 TACCCAGCAC GGTGTTTAAT AAATGTTTGT TGAATGAACG AACTAGTAAA
49651 ATTTTCAAAT CATTAGAGCT GAAGTATCCT TTAAGATTCT TTAGTCCCTC
49701 ATTTTACAGA TAAGGAAGCT AAGGCTCAAG ACATTGTGTG GCTTGGCCAA
49751 AGGCACACAG CAAGCTAAAG GCAGAGGGAG GACAGGACCC GGCTGTCTCA
49801 ACCCCCTGGC TGCTACACTT CCTGCAGCAT TTCTAATTCT TTTACCATTC
49851 TTGCGAGGGA TTTTACAGGC ATGTACTGCT AGAGCCGAAA TAATTAGAAG
49901 CCTCTTACTA CTCATCAGAA AAGCTATGTG AGCCCCTAGG GAGGACACAG
49951 CTAGCCTAGA CTCTGCCTCT TTGCCCTCTG CTGCTTATTA GCAGAATGTA
50001 AGTGGTTGTG TATGATGATT AGTGTAAGTA GGATGGGCAA ATGCACACCT

FIGURE 3, page 13 of 57

50051 TTCCCACCTT CAAACTCAGA AGTTGTAACC AAGAGTCACA CTGACTAAAC
50101 ACTCCAATTT CCCTTTCTGT TTTTCTTAAC ATATGTCCTA TTTTACCAAT
50151 AATAGCCATG GTATATTAGT CATGGTATTT CACGCTAGCT GCAGAAATAA
50201 CTTCCAAATC TCATTGGCTT ACTCAGTGAA AGTTTATTTC TTACTCATAT
50251 AAAGTTGAAT GTCCTGGTCA GGCAGTTATC TAAGCCACAA CTTGGGGATG
50301 GGGATGCAGG CAGCTTCCAT CGTATTGGCT CCACCATTCA GGGATGGCAG
50351 AGTTGCTCTG GCATAATCCA ACCAATAGAG GGGGGAGGTT TGGCACTTGT
50401 CAGTTAACCA CCTAGCCTAG CATTGACACA CACCACTTCT ACATACACTC
50451 CCCTAGTCAT CATTCAGTCA TGTGGCCCAA CCTAGATGCA AAGGCATCTG
50501 GGAAATGTAG CCCCTATCTG GTCAGCAACA ACTTTGCACT TGGAAGGGGA
50551 GCCTGAATCG TTATTGGTCT CCAACACATG TAACTAGCAA TTATACAGAA
50601 CGTTATTTGT CAGGCAATGT GCCAAGAATT ATTTCATTTA ATCTTCACAA
50651 CAATCCTATG AGGTTATTGT CCTCTTTAAC GTATAGATGA AAAAGTTGAT
50701 GGTAGAGATA TAACTTAACT AATGCAAAGT TGCATAAGTG GTTGGTAGCA
50751 AATCCAAAAT TCAGGCTGTT CTCTCCAGAG CTCAGGCTCA TGATTGCTGC
50801 ATTCTACTGC TTTGAGCTTC TGATCTGAGA AAATGCATCA GCCACTAAGT
50851 AGCCTGTGTA GTCTCCAGCA ATTACTTTCC TCCCTCTGGA TCTTGGTTTC
50901 ATTCTCTGCA AAGTGAGGAT GTTTAACTGG ATAAAATCTG ATGTCACCTG
50951 CCAGCTGGGA CATCATATGA TTCTCAGGGT AAGCATATCA GGTGGGTGGG
51001 GTCCCCAGTG ATGCTTGACC ATAGCAAAGC CCTTTCAAAG GTTTCTTAGC
51051 ACACCACATA AATGGAAGCC TCACAGTGTC CATGTAGGAG AAAGCAGGGC
51101 AAAGTATTTC CATTTACCCA ACAAAGAAAT CAACATATAG TAAAAAGAGA
51151 GTGTTTTCCC ACCAAGGCCT CAGATTGACT AGCGGTAGCC TTGGAAATAG
51201 GACTTTATTT TGTATAGTAC TTTTGCCACC AGGGTGGGGG GGAAAAGAGT
51251 GCTTCTTTGC CCCAAATGCT GGTTTCATAA AACCTAAAGA TGTCACATGG
51301 AAACACACCA TTCCCCCAAT CCCCCTCAAA AAACTACTTG CACTTAAATG
51351 AAAGAGTAAA GCTGTAGGAC TTTACTGAGC AGTGTCCTGT GGGGTCCTTG
51401 CACTGCCATG CTCTTGAGGG GCTCGAGGTG TATGAATTCC CCAGCATTAC
51451 TTCTCCTTAG AGGTTTCAGA TGAGCAGTAT GAGCTCCAAA CTCATGCTAG
51501 ACCCAAGTAT TTCATGAAAG AACAATCCTT GAATGACTTT ATACAGCAAA
51551 GCTATATTTC ACTGTGTCCT AGAAAACCAA TTGTGTGTGT TTGTGTGTGT
51601 GTGTACAACT GCTTGTGTTC TTTCTACCTA TGTCCCCCTG ATGCCTCCAC
51651 ACAGAACATC CCAAACTCCA TTTCAGGTTC CTCTTGAGAT TCCCAAACTT
51701 GGAAACAGGA GATGCTTCAA AGGCCTCTTG GAATGTCTTT TGAGGCTTTA
51751 TATTGTGATA TGTTGGACAG ATGGTTAAGA AACAGAAGAA GAGCATCACC
51801 AAAAGGATTT CTCATTTTAT GTGGAGATCT ATTAATATTT GCCACTAGCA
51851 AAGGCATTCT TTCTTGGGAA TGAATTATGC CCCTAGAATC AGATTGACCC
51901 CACAGAAACA AGGGAGAATA AATAGAGACT TGAGCTTAGA CCTTACAACA
51951 TGGCCAGAGC TGAAAAGGCT GAGCTCTAGG CAGAGAAGAT GCAAGAGCAG
52001 CTTCAGAAGA CCTGAGAGCT TATTTGGGTA GGTTCCTCTG GTGTAAAGGG
52051 TTCTTGTTCA CGTTTTCTTC CAGAATAAGA AAAGAACGCA AGGTGTCAGA
52101 GGGTGGATGG AAACAGGGTA TAAAGCAGGA GCATTTGGAA TCTGCCCTTT
52151 GTAGCCTGGC CCAGAGAGCG TCAGGCAGCT TGTTGGGTAA TAAGTAACAC
52201 TGGCATTTTT CCCATGGTTC TGTCATCTTA AAGAGCAGGA TACATAAAGG
52251 GATTCAGATG TCTTGTTGGT TTGGAGAAGC TTCTTTTTAA TACCTTGTTT
52301 TAAAATTTAC CTGGAATTTA TTTTAATCAG GTGTGGTAAG ATGCACAGAC
52351 ATGGAGATGA CAGTCATGAA GGAAGAAGTA TTTATACTCA CAGATCCCTG
52401 TAAATAGGAA GCATGGCCTC CATGCAGGCC AATGGGGAAG CACCAGGGTC
52451 AGCCGCAAGG CAGAAGGAGC AAGAGGAAAA CATGGACAAG AGGCTCTACT
52501 GTGGATTCAG TGGCAAAGAA TGGGAGGGGC AGAGTAAGCA GGTTTAGGAT
52551 TATCGGGTTT GAATGACTTG ATTGAGCTGT AGGGTGTAGA GACTGCCTCT
52601 ACTGTCTGGC ACCAGGGGTA ATTAGGGCAG CTGGATAGTG GTCTGGAGTG
52651 TGAGAGCTCC CTAAAGGAGG TGGTTGGAGG TGTAGGTTTT GGATTGGTTG
52701 ATCTGTATAT GAAAGGTGCA CGTGCAGGTT GAGTCCTCTA CTATCACTAG
52751 AAATTGGCTG GTCCCAGGAG AAGTAGTCTC TCTAGAGACA GCAATGCCCC
52801 AGATGTCAAA GCATCAGAAA ATACAGAAAA AAAATTAAAA GCATGATTAA
52851 TTCATACTCA CAGGTCTAGT TTTTGTGTAG TTAAGAGCAA CCTAAAGAAG
52901 TTGATAACTC GTGTTGCAGG TCAGGTTTCC CAGAAATCAT ATTCTCAGAT
52951 GAAGATTTGC ATGAAGGAGG TTTAATGCTC AAACTAAGCC CTAAGGCTCC
53001 ATACCTGTGG AGGAAGTGAA AGAAGCCCAA CTGGGCACAG AAGGTGGAAC
53051 ACAATGCCAG TCACACAAAG ACCTCAGTGG ATCCTGGGCC ATGAGGAGCT
53101 CTAAGCACAG ATGACCCTTC AGAAATGTCT CCAAGTGGGG AAAGGAATCA
53151 TGCTAGTCAC TGGATGTGGG CTTCCCACTC CACCCCATGA GGGCATGACC
53201 TTAAGTGAGA GAGCTCTTTG GACACAGGGC ATCCTAAGAG GGGCACTCAG
53251 CAGCCACATT GGGCACCAAG ACTCTCAGCA GCTAGAAGAA GAAGGTATAG
53301 TCCCAAAGGG GAATCTGGGC TGCACACCTT AGTATCCATT AGAACTGGAA
53351 GTAGGCTGAA TCCCAGGCAG GGATCCCCTG GAGAACACAG GTAATTTTTT
53401 AAAAAATCAA GCTATGTGTC TGAGGCTATG TGGTAAGACA TCTCAGTTTT
53451 CTGCTAGGAA AAGCCACCAA ACCAGATTGG CTTATTCATG TTGAAAAGTC
53501 TGAGAATCAC ACTCAGATGT TGTTGATAAT TCTGCTTGGA TAAAATTTAT
53551 CTATTGGTAT GCTTGTGATA TAGCAGTACC ATTGCTAAAA ATTCCATGCG
53601 GAGAATCCAA TCTGCATCAT TTTCTTTCTC AATGATTTGT TTTTAAAGGC
53651 AGAGGTTCGG CTGTGCCCCT TTAAACCTTC TGTGCAAGTG CCAGCTTCCT
53701 TTCAAATGGA GAAGCAGCAG CCCTGTCAGA AAGGGTGGCT GGAGCTCCCC
53751 TTTTGTGAGA GGAGGAAAAC TTACTGGGAA TTACCTGTTC GAGAGCCACA
53801 CATGAAGGCA TACCACTGCT TCCTCTGACC TTCCAGCCGG TATATTAATG
53851 ACATACTGTT GTACCTGAGA ACCAATGATG AAGTGGGTGA TGTGCCTGGC

53901 ACCTTAAAGG CCTGGGCCTG CTTTGACAGG GGAGATGATA CACAACATGG
53951 CTGTTAGCCA GCTCTCACTG CATCTGGAAG CACCATGTTC CTTAGAGCCA
54001 AAGTTCTCAA ACTGTGCTTC CTGCTGGGCT CCACAGATCC TTCCCGTTCC
54051 ACCCTGCACA CAAACGTGCA CACACATACA CACACACACA CACACACACA
54101 CACACACAGT GTTCTCAATG CTCGCCATTT AGTTAGTATG CACCAAATAT
54151 GTGTAGTATC TGGTTCCACC CCTGGCCTCT CAGACAATTA TTAGTATTTT
54201 TGGGAGCGGG GAGGAGAGTC AGGAAGACCC AAGCGCCATA TTTATTATTT
54251 CCCCAGCCAC CCCGGCCCAG GCTACATCCA AGTTCAAAGT CTATGACCCC
54301 CTCTCTGAGC TTTCAGCACT ACCTCCCTTT GTGGGGGAGG GGGGTGCCAA
54351 TTCTCTTTCT TCTCATCATC TCCTGTTGCA AAATAAAAGC CTAGGCATTC
54401 CTTTGAGAAA CTTGGGCCTG GCATTGGAAG GCGTCTGACA AAGGCTTTGT
54451 TAAATGAGTG GAGGGAGGGA CGGTCTGGGA GATACTTTTT CAGGTGGCAT
54501 AGGACCTCCG CTTCTTCCCT TCTCACATGA GAAGGAAGAT TTTTCTAGAA
54551 ATCTACAGGT GTTTAAGCTG GAATGTGCCT CAGACATCAT CTGGTTGGAC
54601 CCTTTCATTT TGCAGATCTG AGGCCTAGAA AGATTTGGTA ACTTGCCCCA
54651 GGTCACAGTT GACAGAATTG CTCAGTGAAA AGTCCAGCAT AAATACCCCA
54701 GCCCATGTGG CCACTGGCTG TGTGCTCAGC TAGTGAGGCA CACTTACTTC
54751 TTAATTTGTG CCACCCACTT TTCAGGCTCC CTTAGGACAG CCTCCACCTG
54801 CTCCTACTGT GCTTCCCATC GTCCCTCTCC TCAGGCACAG GCTGAGGAGT
54851 AATAAGAGCA CCTGATATGT GTCAGGCCTT ACTGTGTGCT AGGAATTGTG
54901 CTAAGTACTT CCTATGAATT TTCCATTTAT TCTTTATAAT AACTTTGTAA
54951 AGTTAGAGCC ATTATTCCAG AAGGGAAAAC CGAGGCAATG GGAGTCAAAG
55001 CAAAGAATTT GGGCTTTTAA CCATTACACT ATTTTGCACA AGTAGCCAGT
55051 AATGAAAAGG CTGCTATCCG GAATCATCTT TGCAAAAGGT AATTTCTTTA
55101 GCACTTTATC AGAAGAAGGG GGCTCCTTCC TCAAATTCTG AGGGAAGAGA
55151 AGTGGGGAAG AAAAGATGAC TGAATCCAAA GCTCGGGCAG GGAAAGCACA
55201 TCGAGTGCCA AGTGCGCTGC GCTGGGGTCT AGTCCTGACT CAGCCGCCAT
55251 CTTCCCAAGT GCTTCCTGGA ATTCTCTCCT CTCGTGGGGC CTCAGCTCCT
55301 TCATCTTAGG AAAGAAGGGT AAAGATCTAC AGACAAATTG ATCTTTAAGT
55351 ATCCTTAGAG CACTACCATT TTCAGAATCT AGGATTCTAT ATCCTTCCAA
55401 TTATCTCTGT GTAGGGAATT ATTGGTCGTG TCTCCTGATT AGGGAGCCGG
55451 ACACTCGTCT GTCAGCCCCA CCTGGCTCTG CAAAGTCCCT TGTGTATCTG
55501 CCCTGCCTGG TCACGGGAGA GGAAGAGACA AGGAAACACC ACCGCTCCGA
55551 CTCTGTGGAG CACGCGCTCT CTCCCACCCA CACACCCGCT CAGGAGAGGA
55601 GGAACCTGCA CATTTGAGTC TCCTCAGAGC CTCTGCAGAC TCCCAGCAGG
55651 GGTCTGGCTT TCCTCTCAGG TAGCACAGTC ATGCTGTAAA CTCATTTGGG
55701 TCTTGCTTGG TATGATAATG CGTTTAGTTG AAGGGTTATA TAATTGCAGA
55751 GTCGATGATG ATCTCTAGGC CAATTTAAAG TCAAAGCTAT TTTTAATGGA
55801 ATTGCCAGAG GAGGGCAGGG ATGGGGGCAG GGAGGAGAGA TGGTTAGAGA
55851 GTGCTTTTGA AACCAACCTC CAACAATTTC AGCCATTGCA TTTCCGAACC
55901 TGAATTTTCA GGGCAGAAAT TGGACAATGC CAATTAAATC AGAGCAGGTG
55951 TATGTGAGAG CTGGGTTCAC CTTCTTGCAG CTACAGTTTT ATTTTGAATA
56001 CTGTTGCAGG TAGTGAAAAT ATGACTAGGC TGAATAAGAG ATCTCAGTCT
56051 ATTCCCAGCT CAGCCAAAAG CCCTTAGTGT GTCCTTGATC AAGTTACTTC
56101 CCCTATCCAT TTCCTTACCT GCAAATGAGA AGCTTGAACC AAACTATCCT
56151 AATGTCCCTT TCAACTCTAA AATCCTAGAT GATCCTCAGA TGTCAACAGT
56201 GCTGAAGCCC AGCACTGTAA GATGTCAGGT GGTCCGCAGA GGGTGAGGCT
56251 CTTCCTGCTC AAATTATTTC TTCCACCCAA GACTCCTCAG TTACCTCTGT
56301 ACACAACCTT GCAGGCCCAT CTAAGTATCC AATAACCTGG GGCTTTAGTT
56351 TACAAATTTT CTTGGGGAAG AAGGTAAAAG GGATCTAGCT TTCTGGGTTA
56401 TGAATGCCAT GTAGGGAGGG CATGGTTTGA GTTAGTCCTG GTGCTGGGAG
56451 TTCATGAGAC TTATTCTCAA ATCTTCAGAG AAGAAAATTC CGTGAACACC
56501 TGGGAACATC AGGAAAAAAA AAATGTCCCC TAGGCTACTG TCAGGTTAGG
56551 CTGCTGGTTC TGATTTGACC TTGAACTTGC TATAATTGAA CAAGATAAGC
56601 ATGTGACCTA ATGAAATACT TTAAAACTTG TAGCTTCCTT CAGCACAGAA
56651 GTGGCTCTCT GAACCAATTT TAAGCAATCC TGGCTCTATC TGTGCATGTT
56701 GATTTAGCCT GTGGTTATAG TGTTAACAAT TTAGTGATTC ACCTCATTTT
56751 TAATCTCTCT TTCCCTTTAG CAGGATCATT TTCTCTGTGT TAAGGGATCA
56801 ACATTGAGGT AAGAATGGCT AAATAATAGC ATCTTCTGGA ATACAAATGA
56851 CTTTATAAAT AAAAGAAGAT AAAAGGAAGA AGTAGGATGA TTTCTCAGCT
56901 CTAATACACT TAGCAAATGC CATATGCTTT CTCCTGCGTG TACTGGTCAG
56951 GCCAGTTCTA GATACAATCA TGCGCTGCAT AATGATGTTT TGGTCAACAG
57001 TGGATTGCAT ATGTGACGGT AGTCCTTTAA GATTATAATA CCATATTTTT
57051 GCTGTGCCTT TTCTAGGTCT AGATATGTTT AGATACACAC ATACTTACCA
57101 TTGTGTTCCA ATTGCCTACA GTTTCCAGTA CAGTAACCTG TTGTACAGGT
57151 TTGTAACCTA GGAGCAATAG GCTATACCAT ACAGCCTAGG TGTGTAGTAG
57201 GCTATACCAC TTAGGTCTGG GTAAGTACAC TCTATGATGT TTTCACAGTG
57251 ATGAAACTTC CTAATGACAA ATTTCTCAGA ATGTATCCCA GTTGTTAAGT
57301 GAGGCATGAC AGTACTATAT CTCAAGACTG TCCCCAAGCT GAAGTCTCCA
57351 GTGGACACAA AGACCAATGT ATTTAGTTGA ATCGTGGACC CCAAAAGTTC
57401 AAGTCCACCC AGAACCTCAG AATACAAGTT CAAGTCCACC CAGAACCTCA
57451 GAATACAATT TTATTTAGAA ATAGGGTCTT TGCAAATGTA GTAAGTTAAG
57501 ATGAGGTCAT ACCAGAGTAA AGTGGGCCCT AAATCCAATA TGACTAGCAT
57551 CCTTGTAAGA AAAGGAAAAG GAACACAGAC AGGGGAGAAG GCCATGTGAG
57601 AACAGAGACA AAGACTGGAG TGAGGCATCT ACAAGACAGG GAACACCAAG
57651 GATTGCCAGG AGCCACCAGA AGCTAGGAAG AAGCAAGGAA GCATCCTCTT
57701 CTGGGGCCTT CAGAGACAGG ATGGCCCTGC TGACACCATT GTTTCAAATG

57751 TTTAGCCTTC AGAACTGTGA GACAATAAAT GTATATTGTT TCAAACCATC
57801 CAGTTGGTGG TACTTTGTTA TAGGAAACTA ATACATTCAG GATGGAGAGG
57851 TGTCTGGGAA GCCCATGAGA ACAAATGGAA AGAGCCAGAA GCCCTCAACC
57901 TTGGCTCGTC TACAGCCCAT TTTCTTCATT CCCGCATCCA GGCTTTGAGA
57951 TGACAGGAAG CTGTGAAACC TGTGAATTGT CTCCACCGCA AATCCTGCTC
58001 CCTGGTCCCA CCTAGACTGT CAGGGTTGTG TGGCAAGGCT TTCATGCCTC
58051 TCACTGACTG CCTAGTACGT CCCCTCAATG ACTGGTCCAC ATCTTTCTCA
58101 CCTTTCTCAT GCATGGCCCC AGATCCACCC CAGTGCCTCG TCCTCAAGAG
58151 GTGATTTATT CCGAGACACT GATGAGAGCA CTGTCCTTCC TGTGTCTGAG
58201 GGAAGGCATG TAACTCTTGC TTATCTTCAC CTGTGCTCTA GATCCTGACC
58251 TTCTCTGGCA ACCTCAGGGA CCTTGCACCA TCCATTCTTC TCGCCTAATG
58301 GCGAGACTCA GTCTCTCCCT CTCCCTTTCC ACTCTCCCTT GCCATTCTTA
58351 GTATCTTTCT ACAAGCAGGT CTTCCAAAGT ACTGCTTGAG GTCTGAGTTG
58401 GAGGGAACAT GCCTCTACCC TACTAAAAAG AGAAATTCCT CTGCAGAAGA
58451 CCCAAGCTGA CTGACAAATC CCTTTACTGC AACTGCAGCT CTAGCTCCCA
58501 CCATTTTCCT GTACTTACTC TCCTGCTCAG GTTCCCTGGC ATTGCTGATG
58551 TCTTTCAGCC TTTGTGCCCT GGCCCCTTTC CTCCTCTCCC CTCATCTAGC
58601 ACTACCTGTC AAAATCAGGG ACTTACTTTA AAATTTATCC CAAATTATCA
58651 TTGCCATCAT CTCCACTGTC ACCTTATCAT ATGTTTGAAT AGCGTTTCCA
58701 TTTCCCAAAT GTTTTCGCAT GCACTTTCTC AATTGAGCCT TACGAATCCT
58751 AGAGCTGAGA AGGGTAACAA TTTATGAGTC CTTTGACAAA TGTGGAAACT
58801 GACATCACAG AAAGTAAGTT GCCAGCCGAT ATGTCACTGT CTTCAAACTC
58851 TTCTTTGTAT TTTTATTATC TCCCATTATA TTCTGCCTCT TGTAATGATT
58901 ATTTCTACAT TGGTCATATC TTTCCTTCTG TACTGATCTT CGCTTATGAT
58951 AACAAATAAT AATAGTTTAC CTTTGCATCA CACTTGATGG TTTACAAAAT
59001 GCTTCAAATT CAACATGGCC CTTGATCCTG AAGATATTTA TCACTTAAGA
59051 ATCATTATCG CCATTTTAAA ATACAAATTT ATTACTTGGG CTAAATTTTC
59101 TTATTATAGT TGGGATAGGC CTTCATCCAT AGGGTGAGTG CAGTATTTGT
59151 GGACTGTCAT GGCAGCTTAA ACATTTAGTA CTTGAAAATC TGATGCATTG
59201 ATCATCAGAG AAATGCAAAT CAAAACTACA ATGAGATATT ATTTCACCCC
59251 AGTTAAAATG GCTTTTAGCC AAAAGACAGG CAATAATGAA TGCTGACGAG
59301 GGTGTGAAGA AAACGGAGCT TTCATACACT GTTGGTGAGG ATGTAAATTA
59351 GTACAACCAC CAGGGAAAAC AGTTTGGAGG TTCCTCAAAA AACTAAAAAT
59401 TGAGCTACCG TGTGATCCAC CAATCCCACT GCTGGGTATG TACCCAAAAG
59451 AGAGGAAATC AGTATATGAA AGAGGTATCT GCAGCCGGGC GCGGTGGCTC
59501 ACGCCTGTAA TCCCAGCACT TTGGGAGGCC GAGGCAGGCA GATCATGAGG
59551 TCAGGAGATC GAGACCATCT TGGCTAACAC GGTAAAACCC CGTCTCTACT
59601 AAAAATACAA AAAATTAGCC AGGCGCGGTG GCGGGCACCT GTATTTCCAG
59651 CTACTCGGAA GGCTGAGGCA GGAGAATGGC ATGAACCTGG GAGGCGTAAC
59701 TTTCAGTGAG CCGAGATAGC ACCACTGCAG TCTGGCCTGG GCGAAAGAGC
59751 GAGACTCTGT CTCAAAAAAA AAAAAAAAAA AAAGAAAGAG GTATCTGCAC
59801 TCTCATGTTT GCAGCAGCAC TGTTCACAAT AGCTAAGATT TGGAAGCAAC
59851 CTAAGTGCCC ATCAACAGAT GAATGGATAA AGAAAATGTG GTACATATAT
59901 ACAATGGAGT ACTATTCAAT AAAAAAAAAG AATGAGATCC AGTCATTAGC
59951 AACAACATGG ATGGAACTGG AGATCATTGT GTTAAGTGAA ATAAGCCAGG
60001 CACAGAAAGA AAAACATCTT ATGTTCTTAC TTATTTGTGG GATCTAAAAA
60051 GCAAAACAGT TGAACCTATG GACATAGAGA GTAGAAGGAT GGTTACCAGA
60101 GGCTGGGAAG GGTGGTGGGG GGCTTAGGGG GAGGGTGGGA TGGTTAACTG
60151 GTACAAAAAC AGAAAGAATG AATAAGGCCT ACTATTTGAT AGCACATCAG
60201 GGTGACTATA GTAAATAATA ACGTAGCTGT ACATTTTTAA AAAACTTGAG
60251 TATAACTAAA TTGTTTGCAA CTCAATGGAC AAATGCTTGA GGGGATGAAT
60301 ATGCCATTAT TCATGATGTG CTTATTTCAC ATTGCATGCC TCTGTCAAAA
60351 CATCATATGT ACCCAATAAA TATATACAAC TACTACATAC CCACAAAAAT
60401 TAAAAGTAAA AAAAAAAATT AAGAAAATAA AAGAACAAAA GTAGATGTAT
60451 TCTACATGTC TCCATATTGT AAAACTAGAA CCAGTCAGTT AACTTTAGAG
60501 GAAGGGGATT GTGGACTTGA TATAAAGACA ACTTTATAAT ATGCAGAGCA
60551 GCCTAATCCT ACAATTGTCA AAAAGTATAG TGGATTCTTT ATTTATTTGT
60601 CCATGATATT ATAGAGGTCA TTTCTGCTTT AACAAGTAGG TGGGAGATAG
60651 CTAGGTAGGA TATATTTTGT TCTTATTTTT TATTTTAAAA TATTGGGCTG
60701 TGGCTGGACA TGGTGGCTGA AACCTGTAAT CTCAGCACTT TGGGAGGCTG
60751 AGGCAGGCAG ATCACCTCAG GTTAGGACTT TTCGAGACCA GCTTGGCCAA
60801 TATGGTGAAA CCCCATCCCT ACCAAAAATA CAAAAATTAG CCAGTTGTGG
60851 TGGCATGCAC TGTAGTCTCA GCTCCTTGGG AGGCTGAGGC AGGAGAATTG
60901 CTTGAACATA GGAGGTGGAG GTTGCAGTGA ACTGAGATTA CGCCACTGCA
60951 CTCCAGACTG GGAAACAGAG TGAGACTCTG TTTTATATAT ATATATATAT
61001 ACACACACGT ACATATACAT GTATATATAT ACACATTATT ATTGAAAGCA
61051 GCCAAAGAAA AATAACACAT TATATATAGA GAAAGAGCAA ATGATGAGTG
61101 ACTTTATATG TATATATATG TGTGTGTGTA TATATATAAT GTGTATATAT
61151 ATACATATAT ATATATAGGT TAAGAACCTT CAGCACATGT ATACCTATGT
61201 AACAAACCTG CATGTTCAGC ACATGTATCC CAGAACTTAA AGTGAAAAAA
61251 AAAAAAAAGA ACCTTCTGCA TGCCAGTAAC TGTGCTAAGT GATTAGGATG
61301 CAATGGTAAT AAAAACAAAG TCCCTCTCCT TAAAGAATTT TCTATTTAGA
61351 AGGGAAAACT GGTAAATAAA AAATAAATAT ATAAATTACA ATTTGTGAAA
61401 AGTGCTACAC ATGAAAGAGT GCTGAGACAG ACATCAATGG ATAAACTTTA
61451 GATTGAGAAG GGCTCTGACA AAGCAACATT TAAGGTGCAA CCTGAGAGAA
61501 TAGAAGTTAA ACAGGCAGAT ATTGGTGAAA GAGCAGTCTA GGCAGAGGGA
61551 ACATCATTTG CAAAGGCCCA GGGTAAAGAA GATCCTGGTA AGGAAATGAC

61601 AGTGGAAGAA GGTTAGTGTA GCAGGACTGT GGCTAGGGCG GAGAGGCAGG
61651 GAAGTAGTTT AGAATTTCAA TGCAATAGGA AATATGGAAG ATTGAAGGCA
61701 GTTTTGCATT ATAAAATAAT ATGATTGCTA TTTTAAAGCT ACTTTATCTA
61751 AGGATGGAAG ATTCTTAAAT AAACTTGTGT ATACTTGGAC CACACCACCA
61801 TGAGCAGCAG CTGCTCTAAT TCAGAGCAGT CCTCCTGCCA AACGCTGTGT
61851 GAGACAAAGC TCTGATTCAT AAAGGGGCAT TTTTCTCTGG GAGAAAACCA
61901 GTGATCCATC TGTAGAAGTA CCTGAGTCTA AGGGGAGACG AAGCAGCAAA
61951 AGAAATTGGC TTGTGAGGAC AGGGACATTG TAAGAATGAA AAGAGGAAGG
62001 GAGGTGCTGA GCCCTTTTTC TTTTTTCTTT TTCATTTTTC TTTTTTTTTT
62051 TTTTTGAGAC GGAGTCTTGC TTTGTCGCCC AGGCTGGAGT GCAGTGGCGT
62101 AATCTCAGCT CAGTGCAACC TCCGGCTCCC GGGTTAAAGC GATTCTCCTG
62151 CCTCAGCCTC CCAAGTAGCT GGGACTACAG GCCCTTTTTC TTAATCCACA
62201 ACCTTCAGTT GGATTTTGCA AATGAGTCTG TCTTCACTGT TTCCATTCAG
62251 TGGCTGGAGA CAACTTGGAA GAGAATCTCA GAAATAACTC TGGCTGCTCA
62301 CCCAGTTGTT TGTAAATTTT TATTGAGACT CTACTGTGTG CCAGGCTGTA
62351 CCAGGCACTC AGATATGACA GTGAATGAGA TAGGCAACAT CTTTGCCATT
62401 GGAGAGCCTA CACTGAAGTG GACATGAGGG AGTTGAAAGC AACTCTTATA
62451 GGAAATCATG GTAAAGACGT CCAAGAGAAG AAAGATGAAG GGCAAACACA
62501 TGCACGGATG CCAAACATCT ATCAGAGAGA AAGGAATTTT CAGACCTGAC
62551 CTGAATGATG AAAGGAGGTT TTTGGAAAGG AAAATAGAAG GGAAGGACAA
62601 GGGAAATTAT CTGGGCAGCA ATATTTATCT GCTGTGGTGC TTCACTCTCT
62651 CTCTAATCCT TTTCCACCCC AGCCCCAAAT TTGAAAGGAT TGCAGGGAGC
62701 TCCTGCTGGA GTCATTTCTG GTATTAAAAA TGTACAGAAA GGAAAGCTTT
62751 GGTTCTGAGT TTGCAGGCTT CCCTGTCTTT CATTCCTATT GTAGAAAGCA
62801 GCTTATATAA AAAGATGTGC TGTGTGGCCC TTTGAGCTGC TGTGATTGTG
62851 TTAGGACCCC ACTGGATGGT ATTCGCATGA ATTAATCTAC TGTAGCATCT
62901 CTACAAATCA AGAGGCTGGC TTCTGTTTGA AATGTCCCAA GGCTTTGTGC
62951 ACAGGGCAAG CTAAATGTCT CCCTACAGTG AGACTGAAAA TGCCTTGGGT
63001 GCCCTTGTCG ATAGGATCTG ATATATAGAT GCATGTCTAC AATTGCACAG
63051 TGGCTGCTGG CAACATTTAT TACAATCTGA ATGTGAAATG GCTATTCTGT
63101 TCAAGGATTC TGATAAAAAG TATCAGCCAC AGTAGATGTA TAAGGAGCCT
63151 GGTTTCACTG CAACTGACTA CAGTTATCTG ATTTTTTTTT TCTAGTTCAT
63201 TTTTAGTCTG TGGAGCAAAC AGAGATTTCC TCCCCAAATG ATGTCCTTTC
63251 TCAGTCACCA GGGTGTGGTT ATTTGGTTTT ATGTAGAGGA GATAGAAACC
63301 AATCAGTCTA AATCATATTC TGTTGAAATC AGAACCAAAG GATCCACAAT
63351 CTGGCTCCAA TCTAACTTTC CAGCCTCAAC TCCTACCTGT TCTTTGTTAC
63401 TCTTACCCCT CTAAACCACT TGTGGGATCC TGAACTTGTA ACCTGTGCTC
63451 AGACTGGTGC TTTTGCACTT CTCTGATGGG AAAGATTTCT CTCATCTTTT
63501 ATGATTCAGC TGAAGTTTCA ATGCTTCTGA AATTTTTTCC TGCTCCTGCT
63551 GGAGAGCTTG TTTCTTCTGG ATTCCCATAG GTCAGGTCCT GTGTTTGGCA
63601 TTGGGATACA AAGCCAAGTA ACATAGCATC CATATTCTCA AATCCTCACA
63651 ATTTGGTAGG AATATAGACA AGTAAATACA CCCTGTGCAA CCTTTTGTAA
63701 CAGAGGTATA AAAGGGTATG AAATAAAGAA TTTAATCAAA TCAAATTGAA
63751 TATGGGCTTC AACTCTGAGA TCTTCTTCCA TGATGAGGTT CCCAGTTTAC
63801 TCTAGTGAGG TCATGATTCC ATACTGGCAC TCTTCTAGGC ACATAAGGCT
63851 CTATCCTATT ATTAAATAAA GATTATTACC ATTCTCACTG CAAGCAGCAG
63901 CAACCTGACA CCATCATCAT CATAAAATAA GTAAAACNAG AGTTAATTAA
63951 GTGTGAACTT TCTAAACCAA CATTGTATGA GATAATTACT CATAAAAATG
64001 ATTCTTCACT TTCCAAAGGT GCCTCTAAAT ACTAAGATTT CAGTTACAAT
64051 AAAACTTAGA TCCAATTTAC AGATATTAAA TTTGGTCCAT TTTCCAAGAA
64101 TATTTTCTTT TCTCATAAAA TAAAAAAAGT ATGTGAGAAT ATTAGCACAA
64151 AGGGGTTGCA AAATAAATTT TATTTATCCA GATGTGAGAT AAGAGGCACA
64201 TGCGTCTTTT TTCTTGTTTT ACTGCACTGG TTAGGACCTC TAGTATGTTG
64251 AATAAAAGTG GTAAGAATGG ACATTCTTGC TTTGTTTCCA GTTTGCTTTA
64301 ATATGTTTTC TGTCAGTTTT TCATAGATGC CTTTTATCAG ACTGATTAAT
64351 TCAGTCTATT ATTATTTCAG TATGTTATTC AGTTTATTAT TTCATAATAA
64401 TTTTTTAAAC CATGAATGAG TTTGAATTTT GTCATTCCTT TATGTATCTG
64451 TTGAAATGAT CATATCGTTT TGCTTTCTAA AGCTTCTAAT ATGGTTTAAT
64501 CACATTTATT GATTTTTCAA ATGTGAAGCA AATTTAAATT CATGGCATAA
64551 ATCCTACTTG GTCATCGATG TGTTATCCTT TTTGTATGCT TCTGGGTTCA
64601 ATCTGATACT ATTTGTTAA GTATTTGTGG TGTCTTTTCA TGAGAGATGT
64651 TGGTCTGCAA TTTTTTTTTC TTGTAAGGTT TTTGTAAGGG TTTAAGAAAG
64701 CAAGGTCAGG TAAGCTTCAC AAAGTAAGTC AAGAAGTATT TTCACCTTTA
64751 TCTTCTGAAA GAATTTATGC AACGTTGAAA TTATTTGTTT CAGAGATGGT
64801 CAACAGAATA TACCAGAGAA ACTATTTGGA CTTAGAGCTT CCTTGGGGGA
64851 AGGTTTTTGA TAAATAATGC AATTTCTTTA ATACATAGTA CTTATATTTT
64901 CTATCTTACC TTGTGACAAT TCTGATGAAT TGTGTTTTTC AAGAAGTTTG
64951 CCCATGTCAT CTGAGTTGTT AAACTTACTA CAACAAAGTC TTTGATAATA
65001 TTCCTATATT AGCCTTTGAA TGTCTATAAG ATCTGTCCTG ATGTTCCCTC
65051 TCTCACTTTT TTAAAGAAGT CTTGCTAGAG GTTTACCAAT TTTATTTTGT
65101 TTTATTTTAT TTTATTTTTT CTTATTTGAG ACAGAGTCTC GCTTTGTCGC
65151 CCAGGTTGGA GTGCAGTGGC TCGATCTCGG CTCACTGCAA GCTCTGCCTC
65201 CCAGGTTCAC GCCATTCTCC TGCCTCAGCC TCCCGAGCAG CTGGGACTAC
65251 AGGCACCAGC CACCATGCCC GGCTAATTTT TTGTATTTTT AGTAGAGACG
65301 GGGTTTCACC ACGTTAGCCA GGATGGTCTC GATCTCCTGA CCTTGTGATC
65351 CACCTGCCTC GGCCTCCCAA AGTGCTGGGA TTACAGGCGT GAGCCACCGC
65401 GCCTGGCCGA GGTTTACCAA GTTTATTAAT CTTTTCAAAG GACTACATTT

65451 TGGCTTTGAT AATTTTTCCT ATTTTTTATC TACATTATAC TGATTCCAAT
65501 TCTTATCTTT ATTCTTTTCT TCCTTCTCTT CACTTTGGGT TTAATTTGTT
65551 CATTTTTTTT TCTGGCTTCT TGAGATAGAA GCTGAGATCA TTGATTTTGA
65601 ACCTTTCTTC TTTTCTAAAT AAGTGCATTT AAACTTACAC ATTTCCCTTT
65651 AAGCACTGCC TTAGCTGTAT CTCACAAATT TTGATATTGT CTTTTCATTG
65701 TCTTTTATTC AATATATTCT AATTTTTCTT GTGATTTCTT CTTTGGCCCA
65751 TAGGCTGTTT AGAAATATGT AGTTAGTTTC CAAATATTCG AAGACTTTCA
65801 CAGATACCTT ACTATTATTG ATTTCTAATT TAATTCTGCT ACAATCCAAG
65851 TATATACATT ATAAAGTTTC AGCCTTTTGA AATGTATTAA GAATATTACC
65901 AGAGATAAGA AGATAAGAAT ATTACCAGCG ATAAGTAGGG ATATTTCATA
65951 AATAATAGAC GAATTGATTC ATCAAGAATA TACAACAATC ATAAATGTGT
66001 ATGTGTCTAA TAACAGAGTC TCAAATTATA TGAAACAAAA CTGACAGAAC
66051 TAAAGAGAGA AATGGCCAAT CCCACAATCT TTATCTTTAT CAGGTGATTT
66101 ATCTTGGTGA ACATTCCTTG TGCTCTTGAA AAGAAAGTGT ATTCTGTAGT
66151 CATTGGGTAT AAAATTCTAT ATATGACAAT GAGGTGATTG ATAAAATTAT
66201 TTAGATTGTC TATATCCTAA GTTTTGTAGA ATTATTTCAT GAATTACTAT
66251 GACAAGGATG TTAACAACCT ACAGCTATGA TTGTGGAATT GGCTATTTCT
66301 CTCTTTAGTT CTGTCAGTTT TGTTCCATGT AATTTGAAAC TCTGTTATTA
66351 AACACATACA TTCATGATTG TTGTATCTTC CTGATGAATT GGTTCCGTTA
66401 TTATTTATGC AATGTCCCTA TTTATCTCTG GTCATATTCT TTATCTTGAA
66451 GTCTTTTTAA CTGATATGAA TGTAGCCACT TCATCCTTTT TATGCTTACC
66501 ATTTGCATAG TTTATATTTT TCCATTATCT TATATTCACA CTATTTATCC
66551 CTTTATACTT AAGTCCATGT CTTGTAGACA GTATGCAGTT AATTGTGTCT
66601 TGATTATTTT TACTCCTTTC TGACAATTTC TGCCTTTCCA TATAATATGC
66651 TTATCAATAC AGTTGGAGTT AAATCTACCG TCTTGTTATT TGTCACATCT
66701 CCCATCTTTT GTTGTTGTTC CTCATTTCCT TGTTTATTAC CTTCTTTTCA
66751 GTTATTTTTT TTTTGTATTC CATTTTAATT CCTCAATTGG CTTTATAGCT
66801 ATATATCTTT GTATTATTTT TTATTGTTTG CTCTAGGGAT AGCAATATGT
66851 ATACTTACCA CAGACAATTT AGAAATCATA TTGTACCACT TCACATAAAA
66901 TAGAAGAAGC TTGCAGCAGT CTATGTCCCT TTACACTCCC ATTCTTTGTG
66951 CTATTGTTTC CGTATGTATT ACATCACGTA CATTGTAAAA TCCACAATAG
67001 AGTGTTATAA TCTTTTTCCA AATCCTTGTG TGAATTAAAA ATTTTATGAG
67051 TAGAAAAATA CATATAACAT TTTATTCTTA CCTACATACT TACCAGTTCT
67101 GCTTTCTTTT CATTCTTACC TGTTTCAGTC TTATCTGTAA ACCCGTTTTC
67151 ATTTGGTGTC ATTTCCATTA GCATTTCAGT GCAGAACTTC TAGCAACATA
67201 TTCTCTATTT CCATGTATCT TAAAATATCT TTATTTTGCC TTCGTTTTTG
67251 AAATATATTT TAATTGGACA TAGAAATCTA GGTTGGCAGT TTTCTCTTAT
67301 ACTCTTGGGT TTCATTGTCT TCTGATTTCT GTTGTTTATG AGGAAAAGTC
67351 ATTGATTATT TGCTCTTTCT CTATACACAA TGTATTATTT TTCTTTGGCT
67401 GTTTCAAGAT ATTTTTCTCT TTATCTGTGG TTATCAACAC TTTGATTATG
67451 ATGGCCTAAG TGGTATTATT GTTGTTTGTA TTTATTCCAC TTGGTGTTCC
67501 TTGAGCTTCT AACTTCTGTG AGCTTTTTTT TTCTCAGCGA ATTTGGAAAA
67551 ATTTAAGCCA ATTATTATAT AATTTTTCTT CTCCATTCTT TCTACTCTCT
67601 TTGGAACTCC AGTTGTACAT AGGTTAGACT GCATGACGTT GTCCCATAGA
67651 TCACTAAGAC TCTGTTCATT TTTCAATTTT TTTCTCTATG TTCTTCAGAT
67701 TGGACAATTT ATCTTGATCT CTATTAATGT TCACTTATCC TTTATTATGC
67751 CACCTTCAAT CTGATATTAA GGCCATTCAG ATCTAGAATT TCTATTAGGT
67801 TATTATTTAT AGTATTAATT TCTCTGCTAA GATTTTTTGT CTGTTCATTC
67851 ATTATGACCA CAATATTAGG TTCTTAAACA TATTTTAATA GCTGCTTTCA
67901 AGTCCTTGTC AGTTAATTCC ATCTGAGTCA TCTTGGGGTT ATTTTCTATT
67951 GAGTGATCTT TACCTTATCT GTCGGTCACA TTTTTTTCTG TTTCTTCACA
68001 TGTCTAGTAA TTATTTATTG TTTGCTGTAT ATTGAAATGA AATATTATAA
68051 ACAGTATCAA TTACATTATC TTCCTTTTAA GGGTATTGAG TTTTGTTCTG
68101 GAAGTAGTTA AATTACTAGT AGAACTTTTT GTTCCTGTCA AACTTGATCT
68151 TATTCTTTGT TACAGTGAGC CTATTTTAGT TTTAAAGTTA GTCCTAGGGT
68201 ACAACTCTTG CTCTATTGTA TGCTCCTTAC TTCTATCACA TTTATTTCTA
68251 TTGCCTGAGA TAGTCAATGA GTTCTCACCT GAGCAGGAAC TGCAACATTT
68301 CTTGACATGG TCTTACCTAT GTATTCATCA TTCATCTCTC AGGCCTGTAA
68351 GAAGAGATCT CTGTTGGGTC CTGTGGAATC TTGCTTGCAC TTGGACAGCT
68401 CAGCCTTCAG CCAAAGACTT GCAGGAAAAC CCCATAGAAA CATCTGGGCC
68451 CTCTCAATAT TTGATGTTTA GGAAGCTAAA CGTCAAGTAT AGCCTCCTTT
68501 TCTAGGGACC CTATCTTGTG AATTTCACTC ACCTTAACAA CTCAGAACTC
68551 TTATCTTCTG CCTTCTCAGG GGAGCTAAAC TGTCACTTTC TGTGGGCTCC
68601 ATCTTCCTGC TCCACAATAG GAAAGTATCT GCAGAGAAAA GGCTGGACAA
68651 TTGTGTAGTA ATTGCTTCAC GCATTTCCCT TCTCTCAAAG ATTGTAAGTT
68701 TGCACTGTTT GCTGTTCAAT ACCTGAAAAT GATTTCTACA AATTGTTTTT
68751 CCAGTTTTAT GATTGTTTTC AATGGGAGAT CATTTCTAGT ACCAGTTCCT
68801 CCATCATGGC CAGAGGTACA AGTTCAACTT GGATCATTTT AAAAATACAA
68851 ACTGGGGCAT GTCACTTCCT GCCCCAAACC CCTTGGTAGC TTTCCATTGC
68901 TCTTAGAATA ACTTTGTGAT CTACAACATC TTCTTCAAGG CCCCGCATGA
68951 TACAAATTCT GGCTATTTCT CTAGTTTCTT ATTGCACCAC CTTGTCCCTC
69001 ATCCACCTTT TTTTAGTCT TCTCTCTTTC TTTGAACTTC TACCACCAGG
69051 TTTTTTCACA CGTTCTTCTT TCCCCATTAA CAATGATCCA CCATTCTCTT
69101 TCTTTATCCA CTGTTACTCA TCCTCATAAC TGAAACATCA TTTCCTAAGG
69151 ATGGCCATTC CTGGTTCAGT CAGTCTATAT TTCATCCCCC ATCACATACT
69201 CTTGTTTTAC CCTATATTTT TCCTTCAAAG CACTTATTTA AGTTGTAATT
69251 ATGTGTTGTT TATTTTATGT CTGTCTGCCC TCACAGAATC CACAGTCCAG

FIGURE 3, page 18 of 57

69301 GAGAACAGAA ATCCTGCCTC TTTTATTTAT ACCACATCCA CAGTATTATT
69351 AGTGCCTGTC ACCTAGTAGG TATGCAGTAT GTACCTATTG AATAAATGAA
69401 TTGACTTCTG TCTTTTAGAT CGTCTACTCA TTTTATCATT GATGACAAAC
69451 ATAATACCTT ACATTCGTGT AGTCTTTTTC ACTCCTCAAA GAGGATTTTC
69501 TGCATAGCTC CTCTGAGCCT CACAAAACCC TTTAAGGAAG ATTGTGAATA
69551 TTATCAGATA AAGATTGTGA GACACAGAAA AGCCAGATGA TTTGGCAATG
69601 CTCATAGTAC CAGAGGCAGA AATACAGCTA GAACAGTCTC CTGGCCTCTA
69651 ATCAGGAGTT CTTTCCAGAA CACTGCTTCA TCTTCCATTC TCTTGGGTTC
69701 TTTCTATCCT TACTTTATAG GGCAAAATGT GTGCAAAGTA TAATCCCTCT
69751 TTTGCAATGT GTTTTTAGTT TTTCAGATTG GAATCATGTA GGCTTTTTAT
69801 GCCCTTAATA AATATCAGTG AGCACAAAGG AAGTCCTGTG AGGGCTTATA
69851 ATCATTTTGC TCCCATTAAT TCCAACACTG AGCAGTTTCC CCATTTCCAT
69901 TCTTGGCCTT GTGAAGCTCT TTGCTATCCC TGTTAAAATC TAAAGTTGCT
69951 TGAACCTTCT TATTGCAAAA ATGCATCTTA AACATTCTAA TACCTCTTTT
70001 TTAAAAAACC AATAAAGACT ACGTCAAAAA TCAGCCATCA ATCGAGAAGC
70051 CCTGCAGTCA TTTGTGTGCT GTTGTCCCTA AGTAGAAGTG AATGTGCTGA
70101 GCTCTGCATT CCCCACCTAG CTCCTCTGTG ATCAGGGTGG ACATTCCCAG
70151 GACAACTGGG CCGAGGCTGG AAACACCATC TGAATGTCTG ACCACACAAA
70201 GTTGAGTGGC TGATCCAGGT TTAACCTTGA CCTCATCAGC ACCACCTTCT
70251 AAGCAACACT TTGGCTCAGA AGCCCAGTTA TTTATTCCAA GGGATGATTG
70301 AATGCAGTGC TAGTGTTTCT TCAGGGCTTT TGAACTCATT TATTTATCCA
70351 GTCATTTATA AAAGATGAAG AGGAGAACAA GGTAGGCCAA AGTGGCTTTG
70401 TACTATTAAA GGCTGCTTGA TTTCTAAGTA CATGTTCTTT GCCACCTTTC
70451 TGCCATTCCA CATTCTAGAA GCCATGGGTA AGTCAGCACA GGGATCTTAA
70501 CATGATAACA TTGGTTTTAG GAGGTCTCGT GCATAATGGA CCAGACTTAG
70551 AGCACAATGC TGTAAGGTAG TGATTTAGGT GAGCAGCAGA TTCTGGCTTT
70601 AGGAGTTTAT TATCAGATGC TTTTTAAACG ACTTGTGGCC CAGGATCCCT
70651 GCACCCATGG GAAGCATTGT AGCCTTAGAA CTCTGGGAAT TCTGAATATA
70701 ATTCCTGAAT CAATCGTAAG GATGCATATC TGATGCTTAG TGCAAACCAA
70751 GAGGCAGAAT ATTTGCAGGC AGTGTATCCT TGAAAAACAA ATCTAGGTCA
70801 TTTTCCTGCC ATGCTTCAAG CTTACTTTTC CATCCTTCCT GATGGTAGTA
70851 CTAACTACAT TTGTAGACCA TTTACGTGGT CAACACTGTG CTAAGCTGTT
70901 AGCTTCATTC TCTATGAGAC AGGCACTCTT AGCCCAACTT TACAATTGGG
70951 AAAACTGAGA CTCAATGAGA TAAAGTAAAT TCTTTACAGT CATTATGCTA
71001 GTCCATGAAG GAGCTGCGAT TTGCAACTAA ATCTATCTGA TTCCACAGTC
71051 TTTGCTTTTA ACCAGAGGTT AGCAAACTAC TTCTGTAAAG GGAAGACAGT
71101 AGTTATCTTA ATCTTTGTGG GCAACATAGG GTCTCTGTAA CGTATTCTTC
71151 TTTCTGTCAC AATCTTCTGG AATGTAAAAA ACATTTAAAA TTTACAAACC
71201 TTACAAGAAC AGCTCATGGG CTAAATCGGA CCTGGATTTA GTCTGTGAAT
71251 CATAGTTTGC TGACCCCGCT TTTTAACCAG TATGTACCCT CCTTCTCGGG
71301 ATGTGAAAAA TTAGTGCAAT TGCAATGGAA AATAGCAAGA AAATGGTAAG
71351 GGCCTGGAAG AGGCAGCAGG ATTACATCAG GTGCTATCCC TGCTCTGGTG
71401 AGATGAAACT GGGGATCATT GAACCACCTG GCATTTGTTA AAGAGTTCTG
71451 CTTTCCCTCT GAGATTCTTT CAGGAACCTC ACACCTCTAG CAGCCCGGAG
71501 AACCGTGGGC TGCAAGGAAA TGCCTCCTCA AAGGAGTAGA AAACCTGCAG
71551 GATAGAAATC ATCACATCTG TCTGGCTTTT CTCAACCTTT CTCTTCTGCA
71601 CTTTCTTGGA TATAATCAAA GCACTACCAG GAACTCCAGA GTCGGCACCT
71651 TTTCATTTTT GTGTTTTCAT TTAATTATTT CTCAGCTGCT AAGTGTTTGA
71701 CTGTTTAAGG GACTCTAGTG GTAAATATTT GTCTTTAGCC TGGCAGAAGC
71751 TGTGGTTTCC TTTGATGAGC TCACACGGTG TGGCTTTTAA GATGCTGCTG
71801 ACCAGGACAG CTGACTGTCC CCAGTGGGTG CAGTCCCCAG CAGTGGGCTG
71851 GACCCCTTCC AGAAAGCGCT GCTGGGCCAA GAGGCTTCCT CCAACTTCCC
71901 GCTGCCCCCA TCTAACCAAC ACCTCAGTCT CTTCTCCACC TGCTTCCCTG
71951 CCCTCTTCCT TTCCCTCGCA GACACTTTCT TCTGCCTGGC AAAAGGAATC
72001 TTGTTTCCAT GGAAGCCTCA TTAAATCTGC ATCTTGCTCA GTTTGGGTTT
72051 GATCACGGCT GCCAGAAGTA TTTTTAGCCC ATGCAGTTGC GTAATGAGAT
72101 AGAGATTGGG GAAAGGGGGA GGTGACTGTA TAGGCAGAGG GTTTTTTTAA
72151 AAAAAAGTGA GAAAGAGAAG GAAAACCTCT AAAGAAAAGA GTTTTATGGA
72201 ATTGGAAGAA GGATGGAGCA CCTCTTTTGG GAGCATGAGG CTGGTGTTCT
72251 CTGGTTAGCT CTTCCCACTG GAAGCCCATG GACACTTGCC ATAATACCTG
72301 TCCTGGTCAC ATGTCAGGGG AACCTCTGAT CTCCCTTTCC ATGAGCTTAG
72351 TTGGCCCAGC CAGGGTGACA CTTATGCTAG GGAGTGTGAT TGATGTTGCT
72401 GCTTACAGAT TTCCCCTCCC ACAGACCTGA TGGGGCAGCC AGGATAGTGG
72451 CAGAGAAGAA GACAGAGCAA TAGCAGGAAA GAGAGGACAA CACTAACACA
72501 TTGGAGGTTT ATGTTCAAAG ACGGGATCTA GGGGGTCAGA GAAAGCACAC
72551 CTACCATGTA ATTGGTGCTG GAATCTGATG CCAAGTGCAC CCTTGGCTTC
72601 TGAGGTTCTG AGAACTCTTG CTTGTGCTTT TCAGCCAGAC TATGCCCTCA
72651 CCTGCCCCTG TACTTTAAAG AGCTCTTTAG GCTGGAGTGG TTGTTTGCAT
72701 TGGATTGTTG GAGTGTGTGT GCATGTTGTT GTGTTCTTGT ATTACAAGAC
72751 AAAGAGATTA AAAAAAAACC ACATGCAGCT GTCACAGCTA ATGTTTATTG
72801 AACTTTTACT ATGCCACATG GTGTTTTAAG CATTCTATAT GTGTTAACTC
72851 ATTTTCCCTA ATTCTATGGA CTAGACACTT AAACAGTCTC CATTGTACAA
72901 ACAAGGAAAC TGAGGCACAG AGAGGTTGGG AAACTCATTT GAGGTCCTCC
72951 AGCTAATTAA TAGTGGAGCC AGGTTTTGTA CCCAGACAAC CTGATTTGAG
73001 AATCTGCAGT CCTAGATTAG TAACGTGTTG TTGGCCTGTC ACACATTTTA
73051 AATGACATTC TGTACACAGA ACCATTTATA GTAACTTTGT ATTGTTGAGC
73101 TGAAAGCAGT CTGCAGATGT GCTGCTGGGA TTTCATTCAT CTTCAAAGAG

73151 GTGTTTTTTT TTTTTTTTAA AGGAAAATGC TTTTCTGAGG GTGGTATCTA
73201 AATTCATAAA AATCTTTACG ATCAAGATTT TCACAAATTT CATTCTGACT
73251 CTGTTGCATT GCCCTTCTTC CCATATTCCC AGTTAGTTTG TATTGATTGC
73301 TGCATCTCCC TTGAGCCCAT GGTCCCCCAC AACATTTCTT GCAGAACTGT
73351 GTCCTGCCTT CACACTGTCA GGCAGCAGGA GCCTCTCTAG CGGCCAGCCC
73401 ACAGTCCTGC AGCTCCTTCC TCAGGACGTT TAATTTCCCA CATTTCTATG
73451 CAGTTACCTC ACAGAAGGAT GGCTACGAGG GCCTCACTTG GCTTGGCAAG
73501 TTGGTCCCCT TTTTACTCAC AAGACTCTGT TTATCTCTTT GTTTATCTTT
73551 GTTTATCTCT TTGTTGACCT GCCCCTCTTC AAGGCCTCAG TTTTCTCTGA
73601 AGTTTACAGC TTCCCTCCTC ATCCCGCAAA AGACCAAAGT GGAAAAGATG
73651 AAACCAGAAT CCACTGCAAG CCCCACCTGC CACAGCCTCT CCTCTAAATG
73701 CATTCTCTGT TGTGTTTAGG ACTTGAGAAT GAAGAGGGAC ATGAATTGAG
73751 GATTTGTTTA TTATTCTTTA CAATATCCCT GTGAGCTGAG TACTGTAAAT
73801 ACCCCCATTT GATACATGAG TAAACTGAGG TGTGGAGTGA TAGAGGAATT
73851 TGCTCAAGGT CACATAACTA GTAAGTGGGT GGAGCTGTGA TGTGAAACTG
73901 GGCAGTCTGA TTCTGGGACC TGTGCTCTTA ATCACCAATC TATATTGCCT
73951 CCTACTTGAA AACATCCAGG GAAAATGTTG AGATAGATCA GCTGAAATCT
74001 TCTTGCACAG TAAAGCAGGG GCCACCTGTC CTGGAGTTAC ATTCATCTTG
74051 TTCATTGTCA ACGATTTGTG TTCAGTGACA CCCTCTTCAG CCCAAGAACT
74101 TACCTGGGTG CTGTGACAAT TGGACATGAC TAGGAACAAC CAGTGACATT
74151 GTAGCCCATC CAAACACAGG GTAGGAAGTG GATGCTTGTC ACTCTCTTTT
74201 GGTTATAAGA AGCAGGAACC CAGTAAAGGC ACCTTTTATA TATCTATAAA
74251 GTTGAATATA TAAGATATAT GGGGGCCAGG CACAGTGGCT CACACCTGTA
74301 ATCCGAACAT TTTGGGAGCC CAAAGCAGGT GGATCACCTG AGGTCAGGAG
74351 TTCAAGACCA GCCTGACCAA CATGGTGAAA CCCCATCTTT ACTAAAAATA
74401 CAAAAATTAG CTGGGCGTGG TGGCACACAC CTGTAGTCCC AGCTACTTGG
74451 GAGGCTGAGG CAGGATACTT GCTTGAACCC GGGAGGTGGA GGTTGCAGTG
74501 AGCAGAGATT GCGCCACTGC ACTCCAGCCT GGGTGACAGA GCGAGATTCC
74551 ACCTCAACGA AAAAAAAAAA GAAGATATAT GGGTATGTGT AGAACTCACA
74601 GAAGGGCAAA CAGGCCTTAA CAGGTGCTGA AAACAGGAAC TGGGAAGTTG
74651 CCAGTACCTT CCTGTCTTTT CCCCTGGAAC CAAACGGTTT CTTACTTGCT
74701 TCTCTCTGCA CCTCTGTCTC ATTTCCCTCT CTCTTCAGAT GATTTTTCAT
74751 TGTTGCATCA CACACATAGA AAAATCAGGA TCCACCCTCC CAAGTTTACA
74801 TATCGTTGTT TCAGGCAGCC ATAGTATCCT TAAAACTCCA CATTCCAGGG
74851 AGAAAGCTTG GGTCAAGGAT TCAGCCAAAG GGCAGCGAAA TGGAGTAAAG
74901 ATGCAACTGC CAGGTCTATG GGCAGCAAGG AGGCCGGGAA GGAAGCCGCT
74951 GTTGTGGTCC AAGTGACAAT TCAACAGCTC AAAGCATAAG TAAGTTGTGT
75001 GCTTTTCACA GATGGAGAAA CTGAGGCACA GAAGGAACCT GGCTGGGGTC
75051 CAGGTCTCTG GCCTTTGTGT CAATGCTAGG TCACTGGATG TGGCGTCTGA
75101 TTTCTACAGG AAATGTGGTT TCTCTACTTT GTCCCAGAGC CCACTCAGAG
75151 CACTGGCTGG CCAGGGGGTC CTAGGGCCCT CTTAGGATAG TCTCAGGCCA
75201 ACAGCCCCAG GACAGAAGCA ACCAAAGTGA AGTTATGAAA GAAAGCTCTT
75251 TGCTGATCTG TCAATGGCAC CCTTGTAGAG CCAATACTTA GAACACCTGG
75301 ATTTGAATAC TCATCTCCAA AACCTGTGTT CTTTCTACCA CGTGACAAGC
75351 CCTTGTAAAC CTCACAACGT CTCTATGAGG TGAGCGCTTG CAGATCCACA
75401 CTTTAGATAA GCAAATGGAG GCTCAGAGGG TAAGCAGCTA GTTCAAGGTT
75451 ATGCACCTGA GCCAGGATGT GGACACAGCT CTGTGTCTGA TTCCTAAGGG
75501 CCTGTGCTTT AGCCACTTTG CAATACTGCT GCTGTCTGCT TCATTTCCTC
75551 ATCTGTCAGA TGGGAACGAT AATACTCAAC TCACATGGAT ACTGTATGAG
75601 GAAAAACAGA TAAAAGAAGA GAAAGTGCTT TGAAAACATA AGCAGCCCTG
75651 GCAGATGGGA ATTATTTTTG CTGCTGACAC ACATCCTCAG CCTTGAGGGC
75701 TCTGCTGAGC CATACCCAGC TCAGAGCTCT GGAGGCACCT CCTCCCCATC
75751 AACAGCAGGG GGGACATTCT GTCTTCATCC TGAGCAGGCT GACAAACTGA
75801 ACCCCACTCC TCCCTCAATG TCCCCATGCT GGGAAGGAGT ATAGCTCATG
75851 CTGTGTTCTG TCTTGTTGCT GAGAGAATGC AGAACCCAGA ATTTGGGTCT
75901 CAGCAGGTTG GGGAGAAAAG GAAATGTATT TCTTCCCCCA AGATTTCTTT
75951 TTGAAATATT TTCATTTGTG GAATCAGATT GTGCATGCAA GTTTCTTCCA
76001 GAAATGTAAG ACGTCGTAAT GATGGGAACT GTTGGTTTTA TAATTGAAGG
76051 ATGGGAAAGG AAACTGATAT TTATGGAGCA CCTGTTCTAT ACCAGGCAGC
76101 TACCCAACCA TCAGCCATTG TTGCAATGTT ATGCAAGCTT TATTATCCAC
76151 ATTTCACAGT CTGAGTCTGA CTCAGCAATG TTGTGTTCTA TGTGCTAGTT
76201 CCCACAGGTA GGTGGCTGCA GCGCTGGGAT TTGAACCCAT CTCCAAAGCC
76251 TCCATCTTTC TACCACTGCC TCCCATTGGT GGGGAGGCCA TGGACTGGCT
76301 GTCAGAGATG TCCTTTCCAG TCTAGCAGAC TAGGAAGCTG CTGGAAGCTA
76351 CTTATGCAAA GGTCAGCAAG GAAGGAAACA GAGTCAGAAC TAGATGGGGC
76401 TCCCCTGGCC ACTTTTCCAT GCTGGCCCAC ATGTCCGGCT AGCAGTCAAC
76451 ATTGGGTCTT ATGCAGAGCC ACCTGTGTTC AATGGAAACA TCCTGGACAC
76501 TGCACAAACT AGTGGGAGCC TGTGAGGGAA CAGCCTGTCG GGTTCATTGA
76551 GGTTCAGCCC AACTCATGAG CTAGGGCAGG TACCAGAGGG TGTGTTCCAC
76601 CCAAATGGGG CAGGTAGGCA GGGGACACAG GCTCCATTTT CATGACCAAA
76651 GACTGAGCAG AGAGGCTCTC TGAGCAGTGG CAGAATGGGA AGTGTCAAGA
76701 AGCTTTGTTT GACAATTGAG TCAAGAGGAC AGAAAAGACA GAAAGCAGAC
76751 ATCAGAGTTG GGAAGGCTCA CCCCAGCTCC TTGACAAAGG TGCATGAGGC
76801 CAGTTCTTGA AGCAGTGACC CTGCCTTATG TCATGTGTTT ATCAAAGCCG
76851 GCCCATCAGC CCTGAAGTGG CCTCTGTGTT TAGAAGAGGG CCTGACATGA
76901 TTCTCTGAGA AAGGATTTGA CAACAACAAA GTGTTGCCGT ATGTGTTGTC
76951 TCATCCCCTC AATAGTCCTG TGAGGTATGT GAGACAGGTG TTACTCTCTC

FIGURE 3, page 20 of 57

77001 CACTTGGCAA ATAGGGAAAA GAGGGCCCAG AGAAGTGAAG CTGCTTTCCC
77051 AGGACCACAC AGCTGGTAAA CAGTGTCCAT CTCAGCTGTT CTGTCTCCCA
77101 CACCAAATAC CCTGTGCACC ACGCAAACAC AAAGACAACT GGACAACCAA
77151 GTCATCTAAT GAGTATGCAT GCTATGGTCT CTCTCATTTT GTCTTTCAGG
77201 GCTATACCCT AGGAGAGCTA ATCATTCTTG GTTAGATAAG AAATAGCCAA
77251 CACTTCTGCA GCATGGTAGG CCAAATACCA CCAGAATAAA CTCAGACCCA
77301 AAGAGATGCT CAGAATGTGT GGAGTTAATA CTTCACTATA CAGCTCTAAG
77351 GTATAAGCCT TGTCCATCTG TCACATTATG ACATGTGCTT GCTCCCACCT
77401 CAATTCCTGA TTCCACATTA CAACAAATAC AATTTCAGGC TTTGAACTAA
77451 CAATGCCAAT GTTTCTGAAG CCCATATTAA ATGCCAAAAT CTGAGTCAGC
77501 TACTGGAGGT AGAGACATGA ATAAGATGGT CCATATTATT TTAGAGGATT
77551 CTTTGGTTGC AAAGGGCAGA CACCCAGCTT GAATTCACTT TGGAGAAATT
77601 GGGATTTTTT TGGCTTGCAT AAGCAAAGCA TGAGAAAGAA AGTTCCAGGG
77651 ATGATGAAAA CCAGGAATGC AAATGTCTCC AGAATTCTTT CTTTTTTCTT
77701 TTAGGCCATC TTTTTTCTCT CAAACTGGTT CCCTCCACTG GGCTGGAGAC
77751 GTTACTACCA GCAGCACTCA GACCCACATC TTCAGTTTAA ATGTTGGAAA
77801 TGGACTGTCA GAGAACATTT AGGCCATTCA TTCTGTGGGA GAGATAGGCT
77851 ATGTAAAAAG ATAGCCACTC CCATGTGAAC AATGTGGTTA GGATTAGAGG
77901 CATGAATATA CCCCAAACCA GGGGTGTGGG AAGGAGGTTG ACACTCTAGG
77951 TGATAATACC CAGACCTTAA GGAGCTTTCT GTCTAGAGGG AGGTATGGAC
78001 ATGGACAAGT AATCAACAGC TACAAAGCAG AGCTGCCAGC TCTGCAACAC
78051 AAGAGCCCTG AGAGGCATGA CAGGGGCAGG GTGGGGATCC ATGTGGGTCT
78101 GGATTGAAGT GAGGAGGGGC ATCAGGAAAG CATTCCAGGA GAGCTGAGGG
78151 ACACTTGAGC ACACCCTCAA AGAATGACTG GGGGTCATGA GGTATACAAG
78201 GGAGGAAGTG CACCCGAGAC AGAAACAATC ACATAAGCAA AAATGCAGAA
78251 GAATATGAGG ATCGGGGAAG GGCAAGTAGC TCAGTAGTGT TGGAGGCCAA
78301 GGGACACGAA GGAAGGTGAT AAAGCCCTGA TGTTAAGGAT AGAAAAATCA
78351 AAGTCCTTTG AAAATCATGT GGAGTTAGGA TCTCAAGAAC CCTACAAGGA
78401 TTTCTTTAGA ATAGAATCAA AGAAAAACAA AGTTTACAGT CTGTGAGGGT
78451 TGCATAGGAA GTAACGTGGT GAGAAATGTT GGCTTGAGAA CCACATATCC
78501 ATAACACAAT GGTGTTTTAG AGGATTTGGG GGAAGGGAGA GAAAATCTCA
78551 AATTGTCTCA GTAACTAATG AGCTTTCATG TACATTTAAA ATAGTAATAA
78601 ATGCAATTGT GAGGATGATG GTGAGATGAG CAAAATAATC CAGTTTGTAA
78651 TTGTAGTTAT CAGGCTGGCA TATCCTGCAG GTCACACTTC TAAACATGAC
78701 TTCGAAAAAT CAAAGATCAG CTAAGTTTGA AGTAAGTATT GAAAGAGGGA
78751 GATTATGTTG CCTCAAGTTA AAATAGAACG TAAAAGATGG TGATTCAAAT
78801 GATCAAAAGC ACCAAGCTTC CCTGTTAGGA TTCAAGGGAG GGGTGCGTGG
78851 CTCCGACACC AGATATCTGC AAAGCAATAT GAAATGAGAT CAATAGTAGA
78901 CATTGAAAGA TTGAAACTGA TATAGGATAT TCAAGTACCA GCTTCAAGAA
78951 AATGAAATGA GACCTAATAA AAGAGAGTAG GAGTCAAGGG GGTATACGAT
79001 ATTAAAGAAA GTGAAGAGCC AGGGTTTGTA GGAAGGAAGG GAGAAGAGGC
79051 AAAGAGAGCA GCTCTTTTAA CACAGGAGCT TCCTCCTTTC CCATTCTCCC
79101 TCCTGCTAAA AGCCGAGTTT GTTTTAGCTG AAATGATTGT AAGACAAATT
79151 TTTATTATTA AAAAAGGAGC TATTTTGTGT TGGTTTCCAT TATAAAATCA
79201 GAGCTCTGCT GCCATAAAAT TAAATCCCAT AATAAAATGA GTAGAAAACG
79251 TGATGTCCTG CAGAAAGGAA GATGGCAGCC CACTCAGTGC CATGCTGGGC
79301 TTGACTATAT ACAAGCCGTG CATCTCCTGC TGCGAGTTGT AGCTGCTGCC
79351 CAGCAGTGCA CATTATCGTT GCAGCTGTTT TCCTCACATT CTGAGGTTTA
79401 TGAAATCCCT CATCCATCAA TAATTGATCT TTAGCTCTTA GTCCAGGGGT
79451 TGTCAACTGG CACTCCATGG ACCTTTAGAG GATTGATGGC TAGGTTTTCA
79501 AAGATCTTTG AACCCCCTGA AATTATATAC AAAATACTGT GTGTGAGTAT
79551 GTGCATTTTT CTGGTAAGAA GCACCTGAAT TATCGAAGCA GTTGTGATC
79601 CCCCAAAAAG CTAAGAACTA CTTCCTAGAG CAAAGGGAGA TTTTGCTACA
79651 CTTAGAGATT TACACATTTG ACCAGGGCAG CTCACACAAG TGGGATGCGG
79701 TTTCACATTT CATGGCAGAT CTGCTTCCAG CTATACAAAT TCATCAAGGA
79751 AATATTGTAA TACTTCTATA TGAATCAGGA ATTCACTATA TTTAACTTAT
79801 TTGGAATAAG AACCACTATA TATATACAAG TTTTTCCAAA AGACTGAAGG
79851 TTCTTCCTGT GGCAGGAAGG AATATGATTA GATTCATGAA GCGCCTTTAT
79901 GTTTATATTT CAACTCTGAA AGATAATTGT GACTTTACTA AATCAAACCT
79951 GTATACCACG ATTAGGAAAA TGTGGACTGA TTTGGGGTTC TAGGGGTAAA
80001 ATGTGACCCC TGTGAAGTAC CAATGCACCG TTCTTTTATC TGTGAACGGG
80051 CACTGAGCTT CTGAAATTAA TTAGTAGGCA GGAGGACATG CGCATATGAC
80101 GTGATAGTTT AAGTACTGAT AATTATTCAC TTGGAAGGGA AGAGAATAAA
80151 ATTCAGAACA CAGTATTCCT TAATGGGAAA TCAACTTAGA GGAGGTAGGA
80201 GGGAGATCAA GCAAGAATAT TTCTGGTAAA ACATGCATAA ATCAATGGTC
80251 AGCCAATGTG TTGATCAAAG AAATTATCTT TCGGGGAAAA CAGTAGAAGG
80301 CAATTGAAAA ACAAGCATCA GGCTGCATAA AAACAGCAAA CAAAAGTCAC
80351 AATGGCTTGA TTGTGTGATG AGGTAATTAA TGGCTGCAGT TAGCAAAATA
80401 TGTTCAAAAA AAAGACAGAA AGGGTAGTTA CAGGAGAAAA ACATCCCCGC
80451 AGATCTTCAA AATCAGAAAC AATGAAAATA ATTATTTCAA AAATTAAGAA
80501 AAAAACTCTC TAATTTATAC CTGAATTACC TGGATAATTG GTAAAATTTC
80551 CTGCATATAC AAATCTTGGT CCTCTGCTCC TCTCTCTATA AATAAATAGA
80601 AATGTATGAA TCAATAGTCA GCCAATGTGT TGATCAAAGA AATTATCTTT
80651 TGGGGGAAAA TTGGTAGAAG CCAATTAAAA AACAAGCATC ATATTGCATG
80701 AAAACAGCAA ACGGAAGTCA CAATGGCTCG ACGGTGTAAT GAAGCCACAC
80751 AATATGTATT AAACACATCA TCTACACAGA TGGATTCAAA GATACCTTCT
80801 TTGTGTCTAA GTCCCAAATC TGTGTTTCCT GGCTCTGTTC CCTCATATCT

FIGURE 3, page 21 of 57

80851 AGTCATTCTC CAAGTCAGCA TGCCCAACTT GAAAGTGTCA TTTTCAAAAC
80901 CTGCTTCTTC TCTTCTGGAA GTTCTTCCTC TGCCCATTGC TCCACAATCC
80951 CCACCTCTTT CACCCAGTAG CAAACCTTAA ATTTATCTTT TACTTTGTCT
81001 TACTTCCCCT TCTTATATTC AAAATGTTTC TCACTTGCAT CTCTTTTCAT
81051 TCATTTCATA AGCATTTATG AGCTCCTGTT ATGGTTTGGA AACTGTTCTT
81101 CATGCTGGAG GTGGTCTTAT AAACAAGTAA TTTCAATTGA GTATTTAGTA
81151 TGTTAAGTGC CATCCCAAAG GCAAACACCA GCTGTGGGAG GCTCCCCAAA
81201 TCAGTCTAAG GAAGTTGGGA AAAGCATCTC AGAGAAGATG GTGTCTGAGA
81251 TGGGGAGGAT GTGTGGAACT GGGCAAGGAA GAGAACAAGT AACAACATTC
81301 TAGAAAAAGG CCTCTTTCAG CATGCTAAGA AGTTTGGAGG ACAGAGGAGT
81351 TACCATTCAA AATTTGGAGG GAAGGAAGAG CATACTGAGG TTTGCCACTT
81401 GAACAGATAA TTTCAGCTGT GTTGGGTGAG TGAAGTTGAG TGGGTACAAA
81451 TCAGGTCAGG AATATAAGTT AGGAGACTGT TACTAGAATC CAGGCCAGAG
81501 GTGATGGTGG CCAATATATG AGAGTTTTAG CAGGGAATGA AAAAAAGAAA
81551 ATGTGTTCAT GAGGTAGAAG TAGGTAAAAA CAACAGGATC TGGTTCCTGA
81601 TTGGAAATGG GGGTAGCCTG GAGAGGAAGC CAGAATGCAG GCAAGAATGC
81651 ATAGTGGTAC CATCCACTGA CATAGGGATT AAAGGAGGAG AAGAAGCTTT
81701 GGTAAAGAAA ATAAGAAGTT CAGCTATGGA ATGTTTGAAT TTGATTTCTC
81751 TGATGAGGAG TAGTTCTAGG TGATGATAAT GCTCAGGGTG TAGACTTGAG
81801 AGTGGATGGG TAAAGTAAAG GTTGAGGCTA TTAAAAGGGA AAAGGTCAAG
81851 GAACTGAGGG CCAAGGATTT ATAATAAGTT ATCTTGGGCC ACTAAAGCCA
81901 CGCAGGATGC TGGCAGGAAA CCTATGAGCC AGGTCTTCAA TGTTGAGTCC
81951 AGTGACTCAG GTGTCAGAAG CAGCAGGAGA AGCATTGATA GCCTGATGGG
82001 GAAGGAGCCG TTACCTGAGA GTAGCAGAGA GAGTTATCCT AGCTGACACA
82051 GCTCTCAGGG ATTTGCTTCT AAAGCAATCC TTAGGAAAGA AAGAGCAGTA
82101 TCCACAGGAG ACTGGTGGGC ACTGGCTTCC CCAGAAAACC TACCTAGATG
82151 AATTCTATTC TCAAGGGACT CCTATTTAGA TAAGGGGCTT TGTTAGTTCT
82201 CAGAGCAACA CCAAACAGAT GTATATCTCA TTACTTGCCC CCACAACCTT
82251 TCTGCTCTGG CCACATGGGC CTACCCACTG TCTGCTAAAT GCACTTCATA
82301 TTTTCTTGTT TCAGTGCCTC AGTATTCATA ATCTTCTTTT CCTAATCTCT
82351 GCCCCTCACT TACCTGAATC TTTTGTATTC TCAATGACCT GCTCCATCCC
82401 AGCCCTTTCA AGAACCTTTA ATACCTACCA AGTGAATACT CTCTCCATTG
82451 ATTACACACT TCCTGTAGCA CCTGTTCTAT AATTATGAAA TATTACCTAT
82501 TGTACACATA TATTTCAATC TCTTGGTGGA CAGAGAATCC AATTTATGCC
82551 TTGTCAATTT GTAGCACATT TCCTTGCATA TGTAGATGCA CCATGAATAT
82601 TTAGAGAACT TGTTAGTTAA TTTCCTGTTT AACATGGGCT GCAAAGTTCT
82651 GGTCCATGCA CGTCTTTTAT AAAATAGAAA TGACGGATGG TGCATGGAGC
82701 TTAAATTCCA TGAAGCAGAA ACATATGAGA GATGGAGCTG AATTTGTTTG
82751 CCTGTACAGC TCTTACAGCA ATTGCTTCCA ATTTGTTTGA TTTACCTAAG
82801 AGCTAAAATT GTAAATGGCA GCTCAAATGA TTTTTCTGTA CATTCAGAAA
82851 ATGAGTTTGA ATATTTGTTG GAGAGTAACT GCTTAAGACA TGAAAAAGGG
82901 GGAGATTATA GCTTTTAACT CTTTTTTATG GCAGAGCATT AAGGAAAAAA
82951 AAGTGCAGAT AAATGAGATC AAATGGCAAG TGTCTGAACC TGCTGGACAC
83001 AAGTCCCGGT AGCCATTGAT AGACAGTGTT TATATGACTT CTGGGCCATC
83051 AATAGATAGA TAAGGTACAT CAGCGGCCAA TGTTCCAGGA AGTTTGAGAA
83101 GATAAATGGA AGTTGCACAG CAGCCTAAAA GCTTCCTTAG GAGGGCTGTG
83151 CTCCTCCAGA GCGCCATCTG CCTGTGTCTT CCTGTTCTTC TTCTTCACAT
83201 TAAATGCTTT TCCTTTTCTC ATTTTTATGA TGGTTATCCT AAAGATATGC
83251 TAGCCTGGAC TTTGACAAGG ACATCTGGAG ATAAGAAAGA TTCTGAATTA
83301 TTTTTCCCTT TGGGCAATTG TAGCAATTTT AAAACTATGT TAGATGGCTA
83351 GAGATTCTTG AGAATATTTC TTTTCTTGGA AAATCATAAG GCTTTGGATA
83401 GTGGTACCTA TAGAAGCTGA CATCAGCAGC AGCCTGCCTC CAGTCGATCA
83451 GGGCCTTTGG AACTTCACGG GGCTCCTCTA CTGACAGCCC CATCGGTTTC
83501 CCTCCAGCAC ACGTAACTCA GCATTGACTC TGGGTAGTAG AGGGTGGTTT
83551 ATGGAATCTG ATTCATCTCA GAAAGAGGTG GATGCAAACA CATTCCCAGA
83601 GCAGAAGGCT TGGCATGTCT GGTCTTAGGC AGAGGGAACT GGAGATACTT
83651 GTCCTATTGT TCTTGAGATT CCAGCAAAAA TAGCCCATTA CAGAGGAAGA
83701 AGATATCAGG TCAAATGAAG GCTTTGGTGC TACAACATTG TCTTAGAAAA
83751 AAAAAGAAAG AAATTGGCCA AGTGCAGTGG CTCAGCACTT TGGGAGGCTG
83801 AGGGGGGCAG ACCACTTGAG ATCAGGAGTT CGAGACCAGC CTGGCCAACA
83851 TGGCGAAACT CCGTCTCTAC CAAAAAGTAT TAAAAAATAG CCGAGTGTGG
83901 TGGCGGGCTC CTGTAATCCC AGCTACTCGG GAGGCTGAGG CCGGAGAATC
83951 ACTTGAACCT GGGAGGCGGA GGTTGCAGTG AGCCAAGATC GTGCCATTGC
84001 ACTCCAGCCT GGGCAACAGA GTGAGACTCC ATCTCAAAAA AAAAAAAAAA
84051 GAAAAAAGAA AAAGAAAAAA GAAAAGAAAG AAATTAAATT AAAAAAATTG
84101 TTTTTTAAAC AAAGGAAGGC TTTGGGCTTG GAGTCCAACT AAGCTAGGCT
84151 GGAATCCCGG TTTCATCTCG CTTCTCTGTG CAACTTTGGA TTTTACTGAA
84201 TCTCTCTTAT TCTCAATTCC CTCCTCTGTA AAATGAAGAT AATGCTAGTA
84251 CCTGTCTCAT CAAGTTGAAG GAGACTTAAA TGAGATGTGT TGAAAGCATT
84301 TAGCATAGTA TGTGGCACAT AAAGAACACT CAATAAATGC TGGCTATAAA
84351 GAAGCCAGAG AGAGACTCGG AGGTGATGAG AGAGGCCACA ATTCCCTCCA
84401 TTTCATTGAA AAGCAATTTT TATTATCTCA TTTGAAAGGC AGTATAGTAT
84451 AGTGGTTAAG GACATGCACT ATGGAGCTAG ACCTCCTCAG TTCACTTTCT
84501 GTCTCTATCA TTTATTAGCT GTGACTTAAC CTTCTTGTGC CTCAGTTTTT
84551 ATCATTTTTG AGAGAGGAGT AATAATAGTT CCTACTCTGG TGTGTTGTGG
84601 AGATTTGATG AGTTAATACA TATAAAGCAC ACATAGTAGT GCCTGGAGCA
84651 TATTAAATGA CATGTAAGTA TTAGCTGTTA TTTTATTAAA CAACATGTGG

84701 CATAGGACAT ATTGGAACTT TGAAGTCTTT GAGGCTCTTC CCAGTTTCAT
84751 AAATCAGAGA CTACAGTATA AATATCTGCT TACATGTCTG CTTTCCCCAT
84801 TGGACTGCGA AATCTTGAAA CTGTTTTATT CATCTCTGCA TAGCGTTGGC
84851 ATCGTATTAT GATACCTGAC ATTTACCAGG TGCCAAATGG GACTGGGCAT
84901 GTTGTAGGGA TTCAGTCAAT GTGGGTCATT GCAGGCGGGG AGGTGGGTCG
84951 GGTTAAAGGT AAGAGAAGGG CCTTGGGGCA TCACATTAAG TAGTTACCAG
85001 ATTGAACTGC AAACATTGCT ATCCAGGAGA AATCAGGTCA ATATTTCACC
85051 TTCATGGCAA TACCAGTACA GTCCAAGGAG AATGCATAGA AGGAAAGAAA
85101 TCATAATCTG ATTGTATGTG TTTTTTTAGT AGTAAATAAT AATAATTATT
85151 ACTATTCCTA TACAATTTTG TGTGTTGGTG TGTTTTGTTT TGTTGTGCAT
85201 GAAAAATGGG GTGCTAATCT ATTCCCCTTC CCAACACCAG TGCTCAGAAG
85251 AAATTTCCAC AGATAGAGAA GCTATAGGTT ATGAATTTGG CCTTGATGGA
85301 TTCTGGGTCA CTATTTCTCA ATGTTTGTCC ATGTCATGTG AAGCTCTTAA
85351 GATAAAGAAC AATGTCTTAC TCGTCTTTTT AACTTCTTTA CCCCCTAATG
85401 CCTATCACAT ACTTTGCCCA TGGAAACTCA ATAGACATTT GTAAATGGAA
85451 TTTAATTTCT GAGGTCCAGT AAAGCCTTTT TCCATCCTTC CCCTACTACA
85501 CAGTTTGTCT AACCATGTCT TCCCTTCCAT CATCCACCTT ATAAACGTTA
85551 TTACTCATTC TTCCATCACA TTCTTGACAC CTCCCATGTC CAATGTCAAA
85601 CAAGTACCAT TTGGGAAACA GAATTCTAGG AATCTGGAGA CCTAGAGCTC
85651 TTCAGACCCT GAAATCCAGT TTTCTGAGCT GAGACAGTTT CTTAATTTCT
85701 CACTCCAACT CCGTTTCTCC TCTTTCTCAA TGGATATTTT CCAAGTCTCC
85751 ATTAGGCATA TAGCAATTCC AGAAAACATT CAATTTTCCC TTCTCTTAAT
85801 GCCATGCTCC AAAACACCAC ATTCCCTCTA GACATTGAGC ATTGGAGAGA
85851 GATGGAAAAG TACTTTGAAA ATGTGTGCAT GTGAGAAAAA TGCTAAGTGT
85901 TCTGTCTGGT CACTTCAATG ACAAGTTTGC TACTTTAGAA ACTTGACTAA
85951 ACAGAGTGTG AGGAAAAACA TGAAAAGAAA AAAATGTGTT CAGCTTGGCT
86001 GAATAATGAC CAGCAGGGTG AAAAGATAAG ATAACCACCC GCTCACAGGA
86051 TTTCTATCCT CAAGCCCTAG AAGGTTGACA ACAGCAGACA CTGAAACTAC
86101 TCTTAATGGA GGGTGTGCTA AAGAAGCAAC ATTATAGCCG CTTTTAGGAA
86151 AGCAAATAGG AAAGTTGGTG AAATAGAGAA GATGCCTAAG CATGTGAGAT
86201 ACCACCTCCA TCTTGGAAAA TAACCAAGGT GATACAATGT TATGCAGGAC
86251 CCCTTAATTA AAACAGATTT AGTGATTAAT ATCAGGAGCA TTGTCAAGAA
86301 TCACAACAAC AGCAATTAGT TACTATTGAG CAATTTCTGC TAAGTAATTT
86351 GCAGGAGGGC ATCTCACTTA ATTATCACAT CCTTTTATAG ATGAGAATAT
86401 AGAGGCTTAA AAAGGTGCTT TTCCCAATGT TATTCAGCTA TAAGTGGTCA
86451 GTCATGACTC AAACATAGGT CAACCTGACA ACAAGATCTT CACTCTTAAC
86501 TTCTCTTCTG TGTTGTAATA CCCTTGATCC ATGGAAATGG ACCATCTTCA
86551 TATACTGCTT TTTTGCCTCT GGAATGTCCA GGTATGGATT GGGTAATGCT
86601 CAAAGACAGA GAGGAATAGA GTATTAAAAA GATCCCTGGC CTCATTTTCT
86651 GAAGACATGA GCCTAAGCTG AGCTGTACCA TTTACCATCT ATGTGAACTT
86701 GGGCAGATTT TTTGACACTG CTGGGTCTCA ATTCCTGTAA CTGTCAAGTG
86751 GAAGTGAGCC TAACTGCATA GACTTCACTG GGCTGTTAAG AGAATAAAAT
86801 GAAATAACTG TAAACAGAAG TGCCTAGTGC ACATGCAAAG GATTATTGGG
86851 GCTTTCTACC CTTCAGGGAT TAGAAGTTGA TAGTAGGCAA CAAGTTATAA
86901 GAAATACAGT CAATTGTCTG CTGACCAGGG CTAGAGTTAA TTGTCTCTGG
86951 AAAAAAGGAC TTGCCTCTCT TTCTCTTCTT CCTCCAAAAC TTAAGACGTT
87001 TGCAGCTGAA TCCCCAACAG GATTTTGTTT TCCTTTGGGA GAGAGGAAAC
87051 AGACCAATAT ACCCCCAAAA CTAACCCCAT AATTTCATTT CAGCAGTAAA
87101 GTGAGGTCCT TGATAACTGC CCTGCCCAAC CTGCAGGGTG GTTGGGAAAC
87151 TCTGAATGGT CATGCATGGG GAAGCATTGT GTCCACTGTA AAGAGCTCTC
87201 CGGAGATGAT AAATCTCATC AGAAGGCTTC ATGCTTGAGG CATGGATTCT
87251 TGGAAAAACA ATCACTCTAC GTATGTGGTC AGAATCTAAA GGAGATGCTG
87301 GGGAGAGGAG CTAGGTCAGT CTCCAAAGTG GAACAGTAGA AACTAATCAT
87351 GTGGAGCCTA AACTTATGAA GGTTTTTAAA ATCAGAATTG GCCACCTTCC
87401 TTTGGACCAT GAGCTCAGAT TGTGAGGTGT GACTAGGTCA CGTCTCCTTC
87451 CTGCCCCTGT TTCCCTCCTC TCCCTACCTG TCCCTCCTTG ACCCCAGGAA
87501 AAATTGCCGG GATATGAAAG TTAATTATGA CCCAAGGGAA TTGGTACAGA
87551 TGGGGAAGAA AGAAATGCAT TCAAGAGCAT TTCCATCAGT ATTGAAATTA
87601 CACAGAAGGC TGGTGAATTT GGGCTATCCA TTCTTGCCTC CCTCTGTGCC
87651 CATAATTCCT TGGCCTCCTT CAATTTCATT TTCCCTTTGG TTCAGAGGAA
87701 TGCTTGATGG CTTAAGCTAG CCTCAGTTGG CCAAGCATTG GAGAAACAGA
87751 GAGGTGTATG ACACAGCTAC ACTCCCATGG GGCTTACAGG GCAAGGTGAG
87801 AGAAGACAGA AGTTGTATGT GCTGGGTGCC ACGTGGTAGC TACAAACTAG
87851 AAATGAGACC AGGTTCGGAA GAGGAAGAGG GCTTGCAGAC CTGAGTCATG
87901 GGGACAGTTT CTTCAGGAAA TGGGATCTCA GCTCTGCCTT GTATGCAGGG
87951 CTTACATAAT AAATATGTTT CATTGTTGTT GTTGTTATTG TTGATTTAAT
88001 AAGATTTTGT TTTAAGAAGA TTTTGTAAAA ACAACTGAAC AAATGCAATC
88051 TCCTGCCAGA GCAGGCAGCA GCAAAGGAGA TTAGGAATAT AACCCCCTTG
88101 GAGACGTTCC TTCACCTACC TGGTGCTGGA TTACCTAAAA GCTTCAGCTA
88151 AGTAGGGTCA CCCCCCCAAG AAATTATTTT AAAAAAATTG AAATCTGATA
88201 TTTTTAGAAA ATCTTATCAA GGATATTTAA TTGGACTATT TACACCTATT
88251 TAGGGTCAGT CGGTTTTGGA CAAGTATGCA GGGGTCTTGG AATCAGACCA
88301 CTGGGGTCAA ATCCTAGTTC TGTCACTTCC TAGCTGGGTG ACCTTGGACA
88351 AAGTTACCTG ACTTCTAATA GCTTCAGATT CCTCATGGGC AAAATAGAAA
88401 TGCTACTAGT ACTTAATAGT GCTCTGAGAA GGATTCAATG AGAAGGATTA
88451 AATGTATGTA AAGCACAGTG TTTGCCCATA GGAAGCTGTT ATTTATAAGG
88501 GAGGGGAGCA TCCTAAGGTC CTCCGAATTT AGGAGAACTA AAAATCTTAC

FIGURE 3, page 23 of 57

88551 ACTGACTTCT CCCTTCAACA GCACCTTCAG AATCTCCTTC ATTTTTCATA
88601 CTGTTCTTTC AACCCTTTGA TGAATGAGAA ATTAGGCATT CTTTCCCTGC
88651 AGATTTTCCC AAACCTTCTG CTTTGGCCAA TAAACATATT TTTAGTCCCA
88701 ATCTTGCATG CTCCTTTGGG ACTTTTCATC TGATAAACAT CCCCTCCTGT
88751 GCTCTTGAAT CCAATACCCT TCTTCCCTGC CCTCCACCCA GAGTCTCCTT
88801 GTATCTGCTG TTAGGCACAA TGATGACCCC ACCAAGGTCA GACAATGGCT
88851 GTGGCCTCAC CTGGACCTTG ATGACCCACA TAGCCTAGAG CCCAGAGATC
88901 AGCCACTGAT GGAGGCCCAG AGGGCAGTTG GAAAACTTCA CAAGACAATC
88951 CAGCCTGATT GTTTTGACAT GCCTGACTTC AGGCTGCTAA AAATGAGCTC
89001 GAGGAATCAG ATAGGAAAAA GAGATAGGTG ATGCAATTTT ATTCCATCTC
89051 CCAATTTTCT GAGTCAAGAG TTGTTTGTTT AACTCCAGTT AAATTAGTAT
89101 TTATCCAAAT TTCCTGGGTG CTTGTCCAAA GAAAAGTACC CCAGATCTAC
89151 AAATTAGAAT CTGGGACTGG GACTTAGGAA TTGGCACTTT TACAATTATA
89201 CCAGATGTTT CTAATATGAG TACTTCAACC ACTACCCTTA TAGAAGTGCT
89251 GCCTAGGACC CTCTCTTCTG GCAGGTGAAG TGGAAGGAGG TTTTGTCGAA
89301 GGGAGATTCT CCACTTCAAC TTGAGTGTCT TGGCTTGTAT CCGCTTTGTT
89351 TGGTTCTATT TCACCAAAGG CTTTCATCTT CACATAAATT TTCTTCAGCT
89401 TTAAATAATT AGTTTTGGTA ACCATTGGTA TACTGGAAAG AACATTAGAT
89451 TTGGAGTCCA GGTGGCTTGA GTTCAATTCT CTGCTCTGCC ATTTACCAGC
89501 TGTGTGACAT TGGGCAAGTT GCCAACCTAT CTATGTCATT TCCTCATGTA
89551 AAGATAATCC CACTTCACCA GGCCACTTTT GAGGACCCAG TGAAATGATG
89601 TGTAACCATT TTAGGAACAC TGGATCATTC TACAGTGCAA TTTTTTACAT
89651 CAGCTTGGAG CCTACCATGT AGGCATTCAA ATCCACTGAG TGTATGGAGC
89701 TCCGTGCACA AATAAAAGGA CTTCTCTTTT CTGCCCGTGT ACAACTTTGG
89751 TTTCCTTAAT CAATAGAATC CATGACAATC CTGGGCCATG GTATAAAGAT
89801 GGGACTTTCT TCCTGTGAAG GAGTCTGGTC TGAACATCTT CCAAACTCCA
89851 ACATAACTGA TGTCATTTCT CCACCCAACC CCATTTGCTG TCTCCTGACT
89901 CAATTGCTAG AGAAGCCACT TAAGGAAGGT TCCTGGAGTT AAGGCTGTGT
89951 CTGGGCCAGT GTAGCGAGCA GTTTTCAACA GTCAGTCCTC TTTATCTTCT
90001 CTTTTCCTGC GAGCCTTTAC TAAGCACTGC CTCCTCCTGT CTCCTTACTG
90051 CATCTCCTGA TGGAATGCAC AGGTAAATCT CCTTGGAGAG TACCAGCCAG
90101 GAACAGTCCA CAGCCAAGGC CACCGATCCT CACCGCTGAG CTCCATCTTT
90151 CCTTTCAAGC TGTCCTTCCC CTCCCCTCCC CACCATCACC ATAGCAACAC
90201 AGTGGTATAA AAAAATGAAA GCGCTAAGGC ATCTAAATAT AGTCTGAGTA
90251 TCAACTCTTC CAGCATGGAG CCGAAAACCT AGGGAATGAC AGCTAGAGGC
90301 ATCCAGACGA TAACTGGCAG CCAGGAGGGT GGATAAGTCA AAGGAAGGGG
90351 TCAAGGAAAG AGGGGAAGGA AAGGGAACCA TCACTTGCTG AGCCTGCTGC
90401 CTGTGCTTTC TCATGTCACC CGCACGACAA CCCAATGTGA ATGTTATCAT
90451 CTCCAGGTAA CTGCTGAAGA AACGGAAGCT CAAAGAGGTA AGAGATTTGG
90501 CCAAGGTCAC ACAGCTATAA GCAGTAGAAC TAAGATTTTA ACTCAAGTTT
90551 CTATGGCCCC AGAATTTATG TGTTTCTCTC TCCATACCAC AGGGACAGGT
90601 GCAAGTGAGA GATTTTGCTG GAAGCACTGG GCTTTTTGAG CAGGCCATAT
90651 AAAAATTCTG AGCCCAGAGC TCAACTAAAT TATTGGAAGA GACTGGGCCA
90701 AATATAAGGC TTCTATCTAA GCAGCACCTG TGTTTCTCAA GGACTGAGGA
90751 AAATGAAGGG GGAGGGTTGG CAAGGCTGCA TTTCCCAGGG TGCGTGATTA
90801 TATGGCATGG GGGTGGGGGC CATTATGATG CCCGGACATG GAACTTACAC
90851 CAGTGCAGAA AGGGTGTGAT TAGAAGCCCT AAGCCAGAGA ATGTTCAGTG
90901 TGATAAATGC CATTATTTTT TCCCTCATTC ATTCAATAGA TTTTTTTTTT
90951 AGATGGAGTC TCACTCTGTC GCCCAGGCTG GAGTGCAGTG GCACCATCTC
91001 AGCTCACGGT AACCTCTGCC TCCTGGGTTC AAGCAATTCT TGTGGTCCAG
91051 CTTCCTGAGT AGCTGGGATT ACAGATGTGC ACCACCACGC CTGGCTGATT
91101 TTTTTTTTTT TTTTTTTTTT TGTATTTTTT AGTAGAGACA GGGTTTCACC
91151 ATGTTGGCCA GGCTGGTCTC GAACTCCTGA CCCCAAGTGA TCCACCCACC
91201 TCCACATCCC AAAGTGCTGG GGTTACAGGT GTGAGCTACC GTGCCTAGCC
91251 TCATTCAACA GATATTTTTA TTAAGCATCT GATGTGTGCT TAACTCTGGA
91301 AATATAGGGG TGATTAGAAC AAATGCAGCT CCTGCCCTTG TAGAGCTTAT
91351 TAGGATAGTG GAGAAGACAA ATAAGGAAAC AATTATACAA TTGATTGATT
91401 CTTTACAACT GTAACATGTA CTATAAGTAC ATAACAGAAG AATATCACTT
91451 GCCTGATGAC TTCAGTGAAA GGGAAATACA GAAGTTCTTA CAAATCAAAG
91501 CAATCCCCTG GGCCAATTGT AAAGGTGATG CCCACTTTCA AGGTGGACAG
91551 AGACTGTGCT AGAAGCTTAG CCTCAACCAT GGGTTTATAT GATTGGTAGA
91601 CCCTGCAGAT CCATTCCCAA TGGTGTATCT TCATACTAAT CATGAAATCC
91651 ATCTAATAGC CATACAAGTG AGGTTTTAAA ACCCAACAAA CTAGACTCAA
91701 ATGAAATCTG ATGAGGGAAT TTATGATTTG TTCTTCCTAC AGCCTTTGGT
91751 ATCACTGACA TAAAACTGAA TGTATGTGCT GAGGGTGCTT GTGTCTTGGT
91801 GATAGACAAG GTAGGTGGTC CAGCCCATGG TACTGGCAGC TTAAAGTCAG
91851 CCAGCCATCA GTGGGAAGTG CCTGTGAATT ATGCAGGAGT GGGAGGGGAG
91901 GGAGTAGGCA GTAAAGTAAT GCATTTCTGT GGATCCAAAG CTTTCCAAAC
91951 TACCTGCAAG TCAGCAAATA TGGGGGATGT TGTATGACTA AGTGAGAATC
92001 AGATAATATA ATGTGTATGG AGCTCTTTAG TTCTTCAGAA AAAAATGCTG
92051 TCTAAACAAA TAGTGCTGAT ATCAAAGATA ATGATACAGT ACCCTAATTT
92101 TAATGCTCTG CTACCTACCT GCCAGCTGTT TCCCAGGGAT GTGGTAAAGA
92151 TGAATGGGCA AGATCTGGGA AAGTGTTTTG AAATCCTTGA TTAAAGGCCC
92201 TCCAGGCAGA TGTAGAATTT TAAATGTGTT ATATTACTGC CACTATTGTT
92251 ATGCTTTCTT TTATCACCCC AGAATTTCAC CATCTCCTGT TTCAGGTGAA
92301 CGAGTCTGCC TGACTCTTAC CTGCCCTGAA TGGCATTGGA AAGGTAGCAG
92351 CCCTGAGATG TGCCATATAA ACAAACATGT TTTTAACCAA GGGATCAGGA

92401 GGCCTTCCTG GCTGGCTCCT GTCAGCTGGT CATCACCTCT CTATAACTCT
92451 AGGCTTTCCC AAGCTTATTT TATTTCCATC AATAGGACAG GAATATGTAA
92501 ATGTCCTGCT TGAAATGAGT ATTGGCTACA AGCCATCTGC CTCTGAACAG
92551 AGGTGAAAAG TGGAAATCGG AGGAAGGGCA GATGTCTTTT GCAAGGGAAA
92601 CAGACTGTTT TCTGCCACTG CACTCTGCCC AGGCAAAAGA GTAAAGGAAC
92651 AGCACTCAGG AGAATTCACT GAAGCGAGGG CAGGGTGCAA AAGGAACTTG
92701 AGAAATTGGT ACTGGGACCC AAAATCAGAT TCTGGCATTT CTGGGAAAAG
92751 AAATGGGCAT GGGTGGGGGT TTTATCTGTC AATAAAAGCA TCCAGAATGG
92801 GGCTAGAAGG AAGTAAATTC AGTTGCCACC TCTGCCTACT GGACAGCCAC
92851 GGAGAACTTC TCCTTATCCA AGGTCGAGGA GCCCTCCGGA GTACATACTG
92901 ATACCATTGG TTCTCCCACA CATACCCCCA TGGAGATAAA AACAGGACCC
92951 TGGAAGCCCT GTCCGTGTTT AACCAATGGG ATTGAAACAT GGAAATGAAC
93001 TGCCCCACAA TCCACCCTGT GAGAGACCAA AGAGCAGTGT TGGATTAACA
93051 GGGAATGTTA CCCTGAAAAG GCATTCAGCT TCCACTGGGG CAGCAGGTAC
93101 AGTGCAAAGA TGATCCCACT TAAATTCCTA AGACAGGAAA TAAGGAAAGA
93151 TGTTGTGGAA ACTCAAGACC TCTCAAAGCA TACTCCTTTG TAGTTCTTCC
93201 GCAGACCAGA CCACGGAATT CAGAAAACAC CCTACCTGGT TCCAAACCAG
93251 CACCTGCCAA ACTTCTCACC CTCTTCTGAC CCTGTCCTGG GAGTTAAGAA
93301 AAAAAAAATC ACTTTATTGG TTGCTCCAGT TATAACTTAA ACAGACAGAC
93351 CATCATCAAA TTAAGTGACA TGTACGACTG CTTATTGTAT GCCAGTTACT
93401 GTGCTGTGGG GTTTTGGTTC CATTATCTCA TTTAATCCTC TCAAAAACCC
93451 TGTTAGGTAG GTTTTATTAT TGCACTCATC TTAGATTAAG GAAACTGAGG
93501 CTCATAGAGA TTCGGTAATT TGTCAAAAGC CCTAAAACAT AATTACTGCC
93551 TCCAGATGTC TCTGATTCTA AGGCCCAGGC TCTTAATCAG TAAATGATCA
93601 AATGAATAAT GATTTTCATG GCATCTGTCA TCGGAAAGAA CAATGGAGAA
93651 TATGCTTAAC CAAAGTCATA ACCAAATAAA TGAACTTGAC AGCAGAGCCG
93701 TGATTCTAGC CAAGATGACT ATTTTCATGC ATGTTTTGAA GGCCAGGAAA
93751 AGGAGGTTAG ACTTGTTTGG GAAGGGAAAC AGGAGCTATC AAGGTGAACT
93801 TTTCCTAAGA GTAGCCCAAT AATAGTGCTC GGGAGGGAGT AATGTGTGCA
93851 AGAATAGAGT CAGGGAGACC AGCCAAGTGT GTGCCTCAGC ATCCCTAGCA
93901 CAAATCACAC ACTAAGCATT AAGATTGTCT CTGCAGTGAG AAAGGCCTGG
93951 GACCAAATTT GGGCTCCACC ACTTACTGGT ATTCATTAAT CATTCATGCA
94001 TTCATTCAAC AAATATATAT TGCGTGTGGT CTATGTGCCA GAGACTGTGC
94051 TGGGTGCTGG CAAAGAACAC AGACAAGGTT CCTGCTCTCA TGGAGCTTTT
94101 ATTCTGATGA AGGAAACAGA CCACTTACAG ATAAATAAAT AAACAAGATA
94151 AAGGGAAACA GATATGATGG AGAGTAGCTG GAGGGCCAAG CAGACCGGGC
94201 AGACAAGGTG GTGGCATGTA AGCTAAGACA TTTAAAAAGA ACCTGGTCAT
94251 GAGACTATCT GGAGAAGGAA AGCTCCAGGC AGAGGAAGCA GGTAGTGCAG
94301 AGGCCCTGAG GCAGGAATGA GGACAAGATA TTTGAGAAAA CAGAACAAAG
94351 GCAGGCATGA CCAGGCCGAG TGGGTGGTGG AAAAGTAGTA GAAGGTGAGT
94401 GGGGGAGTGG GGGCATCAAG GTCAGGCTTT GCAGGCTTGA TCAGCGTTCT
94451 CACTGTGGTT CTGGAGCCAG CAGCATCAAT GTTACCTGGG AACTTGTTAG
94501 GAATGCAAAT TCTCAGGCCC CACCCAGACC TGCTGAGTCA CAAACTCTGG
94551 GATGGGGCAC CTCATTGTGT TTTATCGAGC CCTCCAGATG ATTCCGAGTA
94601 TGCTAAAGTT TCAGAATTCC TAGGTTGGAT TATGCAGTTC AATTTTAATT
94651 TTAAATGCAA TGGGAACCTA TGAAAGATTT AAGTAGGGGA GCAGCATGTT
94701 ATAATTTTCT TTAAAAAATT GTTTTTAAGC ACTCCTGCTG AGGAGAGAAT
94751 GGACCATAAC AGGCTAAGAG AAATGGAAGC AGGGAGATAA ATTAGGTGGT
94801 TATTGCAAGA GGCCAGGTAA GAAGAGAAAG TGGTTTAAGT AGGGTGGTGT
94851 GGCAGAGAAG ACGGTTCCAA GCAGAGGGGG ACCACGCTGA CAAATAAGCG
94901 CGGGCCACTC ACGCAAGCCC AACAAGGCAG AAGGCAGAAG GCAAAAGTGA
94951 AGGCCAGAGA AAACTGGACA CCACCTTTCC AGAGCACAGT TCAAAGGCAA
95001 TGTCCTCAAA GAAGACACTC CACCCTCCTC CCATTTCCTC CCTATTGCCT
95051 AAAAATAAGA AGGATACGCG GCCTATGGCA AACCTTGGGC AGGCACGTGG
95101 GAGCTGAGCT CTTGCAAAGG GCAGATAGTT CCTCTGGTGA GAGAGAAAAG
95151 GAAGGGCCAG TGAGGAGTGA AGGAAGAGAC GAACAGAGAG CCCGAAAGGC
95201 TGAGAACGTT GTCTGGCTTC CTGAAAGGCT TAAGGGGTTA GCTCTGGAGG
95251 GTGAACTAAA AGCCCTAGTT ATATTAAACA CACACGCACA CACGCACGCA
95301 CACACATGCG CGCACACACA CACACACATA CACACAGTTG AAGGAGACCT
95351 GCAGTTTCCA AAAACAAGAG TTGTATTTTT TTTGTTCATA TCATGACCCA
95401 TAACAATCTC AAAAGAGAAA CAATCTCTTG TCTTCCTTGT TTAGGCTTAG
95451 GAGAACCTGT AGTAAGTAAG CAGCAGCAGC GGAACTCAAA CTCGACTCTT
95501 CCTACTGTCA TTCTCTCTAT TACACCACAA GGCATCAGAG GACCACTAGA
95551 GTCGCCTCCC TAGGGTTAGG GTTAGGGCAA GGTAAATGAA GTGAGTCAGC
95601 AAGGGCAGGA TAGGAACCTG TCTTTATTAA CATTTTGATA TTTTGTTTAT
95651 CATGGATTTG TTGCATTAAT TGCAACTTTT AAAAATCATT GCATTAAAAT
95701 ATTATTGATC TTGATTACTG AGTTTTTAGG TGTACCCTTA AATGTTGCAC
95751 CTCTGACTTA CTAGTCTCAC CCTGATCCCT GTCCTGGATC TATGCCTGTC
95801 TGTTCTATAT CAGCCTCTTG CTTTGACCAT AAGAATAACT TCAGACCTTT
95851 AAGCATAGAG GAAATAGGAT TTCTGTCTCC CTTCCCCACC TTTGTGATAA
95901 TCTCAGCTTC TGCTTTTAAA GTCTATCTCC CAAGTAGTTT GCCTACTATG
95951 TTCCTCCCAA GGTCACTAGG TTCTGTGAAA CTAGCAGCAG GCTAGATTGT
96001 CACATTAGCA CAAAGGATCC ACTATTCCTG CAGCCGAGCT GGGACAAGCA
96051 CTTAGGCCCA CTGACTCCAA CCCTTCAATA GCCTGGGACC TACGTTGTCT
96101 CCAGGTGGTA TAAAACAAGA ATTTCCCCTT TGACTGGGAG AAAAAGGGAA
96151 GAACTCTAAA TTGGAAAACA GGTCATCTCG AATTCTCACA GGTGGAAATT
96201 TCTGACAACC CCTTTGGGAC CCACAATTCA ACACACCCCA AATGGGGACA

FIGURE 3, page 25 of 57

96251 GTAGCTAACA TGCAACCTGT AGGCTGTTCT GTCATCCAGT GCCACTGTGC
96301 TGCACACCAC CAGGGGGCAG CATTCTCATT GGCTTCTATG TGCCTGGAGC
96351 CCAGTGCAGT TGTGCAACAC TGCAGCTTTG CTTTAGTGTA GTCCCTGATG
96401 GGTTCAGTCA AGAAAATGTC TATAGAATCA GCTAATCTCC CATGCAGTTA
96451 AGTCTCTAAT TGAAATATTT TCTCTGCTCA GCCCAGGGAC AGCAATCTTT
96501 CCTGGATTTG CTATTTACAA GGATCTCTAG AAATTATCCA CCAGAAATAT
96551 GGGCTTTCTC AGAGCTTGAG TGGACAGGGA ATTAAGGTGG AAGGCAGGGC
96601 GTTTTGACTG CATTTGACCC AAGTCCTGAA GAGCCAGCTC CTCTCTCTTC
96651 CTAATTATTA GAAGGTTTTG TTTGGACCCA GTGTTTCACG TGTATACAAT
96701 ACAAACTTCT CTCTTTTCTA CTTGGATCAA ATTTGTTCTC TCAAAATAAG
96751 ATTCCCAGCA GTGAGAGAAG ACAAGACAGA GAGATCCAAC ATCTCTAAAG
96801 CCATGAATCA GATAACCAGC CACTTGTTCT CTTCAGTGCT GGGAACAGAT
96851 ACACTGTTAA ATAAAATGAT TTTATAGATT CTTCTCACTG CCTTTCCAAG
96901 AAGGGGATTT ATCAACTTCA GGGCACAGCA ATCATTTATT CCCAGACTAC
96951 TGGCATGCAT ATATATATAT ATTTACTTCT CTTGACTTAG AAAAAAGAGA
97001 GAATTGGAGT TGTGAATATT CCTGTCTCCC TCACCCCAGC CCCCTTGAAG
97051 TGAGTCAGGA CAAACTTGGG GCCCAAATGG AGCTGTAAGT AACTGAGTCA
97101 CATGCAGAGA TGAAACCTTC ACAGACCCAC TGATATGGAG GTTGAAGATT
97151 AAATTCCCCT TTGAGAATAA CTGGGTAACA CTCATACAGA GACTACTTTC
97201 AAGAAGGCCA GATCCTCCCT CTAATGTATA GTGCAACGTT CCTAACCCTC
97251 AGCCCACTCC GTCATACCCC CACTCACATG AATACACACA TAAGCAGTAA
97301 TATAAAGCAC TTCCCACCAT AGGGCAGCAA AGAAGGAGGG AAATCTTTAT
97351 TATGGAAGAG TGGAAGGAAG GAAGGGAAGG GAAGGGAAGG GAAGGGTAAG
97401 AGGAAGAATT CTCAGGGTGA GCAGAGGAAT GACATGTTTG GGGCATAATG
97451 AAGATAATTG AAGTGCAGAG TTTGTATGGA AAAATTTGAA AATATCAGGT
97501 GGCAGGCCAG GCATGGTAGC TCATGCCTGT AATCCCAGCA CTTTGGGAGG
97551 CCAAAGCAGG CGGATCACCT GAGGTCACGA GTTTGAGACT AGCCGGGCCA
97601 ACATGGCAAA ACCCCATCTC GACTAAAAAT ACAAAAATTA GCTGGGTTTA
97651 GTGGCGCATG CCTGTAATCC CAGCTACTCG GGAGGCTGAG GCAGGAGAAT
97701 CATTTGAGCC TGGGAGGCAA AGGTTGCAGT GAGTCGAGAT CATGCTACTA
97751 CACTTCAGCC TGGGTGAGAG AGCTTTCTTT TTTTTCTCTC ACAAAAAAAG
97801 AAAAGTTCAG GTTGCAGAGA TGGATGGATG GATGGATGGA TGGATGGATG
97851 GACGGATAGA TAGACATTAC AGAGAGTTTC CAATTCTTAG GATGAATTGG
97901 AATCCTTAAG TCTTTATTCT GTAAGAAAGG AAGGGGAGAA TAAAATTTTG
97951 TGATTTTAAA ATATTTTCTA CCCTGTAGAG CTACCCTACA AGGCATGAAA
98001 ACCTTAAAAA AAAAGGCATC TACTTTAAAA GAATAATGTC TAAAAAATTA
98051 GAAATTCCCT CTTTTTGCCC TGACCTTTGG GAAACAGAGT GAGTGATCCT
98101 TTTGAGGTTT TTGGCACTGC CTTGCCTGTG ATCATATCCT GAACCCTAGG
98151 TCCATAATCA TGCAGTTACC TCAGATGTCC CTTTCCCTCT AGCCACAGGT
98201 AACACGCTCT CCAGGCACTG GGAAAGTGGG TAATTAGGAA AGCAGAGGAG
98251 TACCCATGGG CTGTGATGCC CAGTTATAAA CCCAGACATT TCAGAATTAA
98301 CAGAATGAGC ATCAAGTCCT CAAATGGGTC TACATCCATA AACATGTCCA
98351 GCAGTCAGCT CTTTACTGTC AGTAGAGACA AAATGTTCCT ACACTTTCCC
98401 TAGGGGAAGC CACATCCTCA GTAGGTTATC TCTGATGAGT CCAGCTAGTC
98451 ACAGGTATGT AGAAGCTGCA TGCAGCAGAG GGCTCAAAGG AGGGTCCAGA
98501 ATAGATACCA AAGCAAAAGG GGAGTCTGTG CACGTTCTCA CACGCACCCC
98551 GAAACACTCT TTTTGTTCAC AAAATAGATG GTGTAGGGTA GTTCCAAGAG
98601 ATCATTTAGC TCAGGTTCCT GCCTCCATAA AATAAATAAG CCTTCCATAT
98651 TAGTTGTCTG TTGCTGTGTA GCAAATTGTC AGAAACGTAG AGGCTTAAAG
98701 CAATACCCAT TTATTATCTC GCAAGTTCTG TATCTCAGAA GTCCAGGCAG
98751 GCTTGACTGG GTTCTCTGTC CAAGTTCTCG TGAGACTGAA ATCAAGGTGT
98801 TGGCCAGGCT GGGATCTTAT CTGGAGGCTC TGAGGACATA TACGCTTCCA
98851 ACCTTATTCA GGCCATCAGC AGAATCCCGT CTCTTGTGGC TTGAGGTTGG
98901 AGGTCCCCGT TTCCTTGCTG GCTGTCATCC AGGGACCACT CTTTGCACCT
98951 ACAGGCTGCC TATGTTCCTA TTCACAAGAC ACCGTTCATC TTCAAACCAA
99001 AGCAGCATGT AGAATCTTTC TTGTGGCTCG TATCTTTCTG GCTTTCCCTT
99051 CTTCTTTAGC CAGAGAAAGT TCTTTGCTTT TAAGCGTTCA TGCGATTCAA
99101 TCAGGCCCAC CTGGATAATG TCCCTATTTT AAAGGTAACT GTGATACCGT
99151 ATAACATTTC AGGAGTGATA ACAGCACATT TACAGGTTCC AAGGATTGGG
99201 GCAGAACATC TTTGGGGGAA CATTTTAGAA ACTCTGCCTC CCCACTCACC
99251 CATAATCCTT TTAAAAACCA AATCTTGAAG CCTTTTTTTC CCAAAGGCCT
99301 TTTTGAATAA GCACATTTAT ACCTAACTTC ATCAGACACC CACTTTGAGC
99351 AAACACTAGC ATGTGGCAAA ATAGGCTGTA AATCAATCAG AACTATTCTT
99401 TCCCACCACA ATCTTTCTCA AACACATTGG GAGAATCTGA CACTGTCAGT
99451 GGTATACCAG AGCAGACTCC TACCATCTCA CAAGAGCTGA CTGTTAAATG
99501 TTTAGTAATT GTGGACATTG GTTGTTAAAC TATTAGTAGC CTGAAATTGA
99551 CTATAGTGAG AGTATTTTCA CCATGGAAAG CAACCGTTCC AAATCAGGGT
99601 TTCTCTTTAT TCCTGGGAAG CTGGTTTATT AGCTCACCAC TGGCTGTAGT
99651 CCTTTAGGGG TCATTACTTG ACCTCCTGTA GCATGCAGGA ATCCTCTCCA
99701 TGGCCTTTTT TATGCATGGA CATCATCCTA TTTTTTAATA CCAGGAATGG
99751 GGTGATCACT CTCTTATAAG CTAGTTCATC TCCCTGATGG AATGGTATGT
99801 GGTAGAGTTG AAACCCACCT CCCTGGAACT TCCCACCAAC TTCCTTTGGA
99851 AGCAGCACTT GTGACAGCCC CAGAACCATT TGGAGTAAGT AGCATTTCCT
99901 CCAGGAGACA TCTCTCCTCT GGATCCACAA ATCAATAGTT AGATGCAAAA
99951 TCTTTAGAGC CACACTGTTT GAATTCAATT CCCAGCTCTG CCACTTATTT
100001 AGTTATAACC TTAGGCAAGT CTCTTAACTT TTCTGGTCCT CTGGTTCTTC
100051 ATGTGTGGGA ATGGGGATAA AAATAGCACC TACCTCATAG GTTATTATGA

100101 ATATTAAATG AGATAATGTG TGCAGAGAAA ATAGCACCTG GTCTGGCCTC
100151 TACCTATCTA ACAGGTTAGT TGTGAGGATT AAATTACTTA ATATAAGCAA
100201 AATGCTTAGA GCTCTGCCTA GCACAAAATA AGCACTATGT AACTATTGGT
100251 AAGTTAATTT GAAATGTGGT TTCTAGATCT CTCATCATCC TAGTCACCCT
100301 ACTCTGGATG TACTCCAAAG TCCCTCTCAA GATATAGTGT CAGAATTGAC
100351 CTAATTAGTC CAGCATTTGA CTGAAACGCT AGACTTTGAC TCCAGCCCCC
100401 CATCCTTGAC TGGCACTAGC ATTCAAGCCG CTTCTCCTCT TTCCCTGGGT
100451 CTTTAATAGA GTCAGAGCGA CTTCTCCAGG GGATCTTTTG GCCATGGACC
100501 AGTAGCATCC ACACACGCTG GGGCCTTGTT AAAAAGGCAG GCTCTCAGGC
100551 CCCACCCCAG ATCTACTGAA TCAGAATCCA CACATTAACA AGATGCTTGG
100601 GTGATTCATG TGCACATTAA AGTTTGAGAA GCACCGCTTT CAGGGACGAG
100651 ATGACACACT TATTTTAAAG AGAACGCCAA TTAGAGACCC TAAGCCTTCT
100701 CATGGAACAG GGGCCTTCCC CTCAGACCTT GGGAGAGGGG TCAGGGAAAT
100751 ATCAGTGTTG GGTTGTTGGT GACAGGTGGC GGTGGGGGGT TCAGTCCACG
100801 TTCAAAGAGC CAGAAACCTG GCAGGGGAAG AGATGGGGCA GTGACACCCA
100851 ACCGGAAAAA TAAAGGAAAC TACAAGAAGA ACCCAGCTAA GAGATGTGAG
100901 GCTTCTGAAA GCTCCCATGG AAAGGTTCGC AGCTCCTCCA CCTGCTCGGT
100951 CCAGCTGCCC CAGGTCAAGG AAGCTCTGTG AGTGTTAGCT GACCCGGAGC
101001 AGCAAGGATA CATTCAGAAG TGATGAAAGG GAACGCTTCT TGACAGGGTA
101051 AAGAGTCATT CAGTAGGAAT GAGACAGGAA GAGGTCACAG AGTCAGAAGC
101101 CCAGCCTGTA CTCAGAGATT ATTTCTGGCA TGGGAGGGCC GAAGGGTTAG
101151 GAGGCCACCT ACTCACAATA CAATACAGAG GCAGATCCAC TTATTACCTG
101201 CCTGTGCTGC TGGGATTTCA GTGTGGAAAT TCTGTGCCTC CTCACTGTGG
101251 CTGCAGCTTG GGAATGACAT CCAGAGCTTA CCCACCTGCA TAAGAAATAA
101301 GCTATAGGTG TAATAGGGGG ACATAGGCTA AAATCCTAGC TCAGCTGCTT
101351 AATAGCTGTG CGACTGAGCA AGTTACTTAA CCTCTTTGAG CATCTGTTTT
101401 CTCATCTTTA AAATGGAAGT AATCATAATT GACCAGGCCC AGTGGCTCAC
101451 ACCTATAATC CCAGCACCTT GGAAGGCCGA GGCCAGTGGA TTGCTTGAGC
101501 CCAAGAGTTT GAGACCAGCA TGGTGACACC TCGTCTCTAG AAAAAATACA
101551 AAAATTAGCC AGGCATGGTG GCAGGTGCCT GTAGTCTTAG CTACTCGGTA
101601 GGCTGAGGTG GGAAGATTAT ATGAGCCCGG GAGGTTGAGG CTGTGGTGAG
101651 CCAGATTGTG CCACTGCAAT CTAGCCTGGA GACAGAGTGA GACTGTGTCT
101701 CAAAAATAAA TAAATAAAAT AATAATATCT ATGTTAATAA AGCAGAAATA
101751 AGAATGAAAT AAGAGGCCTG ACATGGTGAC TTATGCCTGT AATCCCAGCA
101801 CTTTGGGAGG TCAAGGTGAG AGGATCACTT GAGCCCAGGA GTTCAAGATC
101851 AGCCTGGGCA ACTTAGTGAG GTCCCATCTC TACCAATAAT AATTTTTTAA
101901 AAATTAGCTG GGCATGGTGG CATGCACCCG TGGCCCCAGC TACTCAAGAG
101951 GCTGAGGCAG GAGGACGGCC TGAGCACAGG AGTTGAGGCT GCAGTGAGTC
102001 ATGATCACAC CACTGCACTC CGGCCTGGGT GACAGAGTGA GACCCTGTCT
102051 CAATAAATAA ATAAGAAGAA TGAAACAAGA AAGTTCTTCT TATGGTTCTC
102101 ATGGTGGTGA GCACAATGTA AGCATATATA TTATCTTAGA ATTCTTCCTT
102151 CCTGTATAAA GAAGGCCTCC TCCAATGTAT TAATCATCTG TTCAACTAAT
102201 AAATGCTGCT TACTCCCACT TTCACTCTAA AGGAACTCAA TGGCTAAAGA
102251 GAACCCTTCC CCTTTGCAGC ACCCTGAGGA TCAGAGGCCT GATTTGAATG
102301 TCCTCGATGC AAAGGACTAT TTCAAAAGGC CAGCCAGGCA GCCCAGACAT
102351 GTATTTCCTA ATCGTCTCCA GGTTGTTTGA TAGAAGATCT CCTGGGAGCA
102401 GGTTTCCGCA GCAGCTCAGC CAGGTCTGTT CTGGGAACGC TGTGTGCATT
102451 GGCACCTCCC TTGGCAGAAA GCTTGGAGGA AAGGCAGGTG CAGGTCCTGG
102501 AGCCTCTGAC AGCATTACTG GCTCTAGGAG TAGCTGCTCA GGATAATCTG
102551 TCCCCATGAC CATTAAGTAA CTGCCACTGT GCGGGAAGAA GAACTGGAAA
102601 TGGGGGGCCC AAAAAAATCT GAAAACCCTC ACTTGAACCA GTAAGTTATA
102651 CCCTGGGTTG CTGTTGGAGA GAGCTTCCTT GGAGTAGACA AATGTGGTAT
102701 GTTAAGTAAA CTGGGGATCT AGGTTTGATG ATACTGGGTC TGCAGCTTCT
102751 TTGTCCCACT GAAAATCCTC GGGCATTCCA TGAAAGTAGC CTTCAAAATA
102801 TTTTTGTCTC TAATGACATA TTTTTGCTGC AAAAAGATGA GTGGATTCAT
102851 TTTACGAAGT CTCAAGTGTG TTAGAAATTC ACCATGAGTC ACTCAGCAAG
102901 TTATGTTTGA GGGCGTTCTG TATGCCAGGC ACTGTGCTGG GCACTGGGAC
102951 TACTGTAGCA AGTCAGATAG ACAAGAACTT GCTTGATCTT GGAAGTAAGC
103001 AGGGTGGGGT CTGGTTAGTC CTTGAATTGG AGACTGCCTG GAGATACTGG
103051 ATGCTGCAAG CTTTTGAAAA AAGACAAGTT CTCTGTACTT GCAGAGCTTA
103101 CATCCAGTAA CTAACTAACT AACTTCAGGC TGTGTTGAGT GACTGAAAGT
103151 GGTGGAGCCA GGAGTCCTCT AGATAAGGTA GCCATGGAAG GCCTCTCCGA
103201 AGAGGTGATA AGTTTACTCA GAGACGCAAA CGATCAGGAT AAGCACAGAC
103251 CCCGGTGAAG AGCGTCCCAG GCAGAGGGGA TAGCAAGGGG ATTGCCCTTA
103301 GGTGGGAAAG GGCTTGATTT GAGGACTGGG AAGACCAGTG TGTCTAGGAC
103351 ACATAAGCAA GGGGAGGACG TTATGAACGA GGTCTGAGGG GTCAGCAGCG
103401 ACTGGATCAT GCAAGCTCCC ATAGGCCATG GTAAGGGCTC TGTGTGTACT
103451 ACAATTACAG GATGCATGAT AGGACCTGGG CTGCATTTTT AATAGTTAAC
103501 CCTGGCTATA ATGTGGGGAA GGGATTGAAG AAAGAGGGCA AAGGCAGGAA
103551 CAGGAAAATC TCTTAGGAGG CTACTGCAAA GCCCAAGGGA GAGGTGATGG
103601 TGTTTTGTTG TTGTTGTTGT TTGTTTTGTT TTGCTTTGAG AAGGAGTCTC
103651 ACTCTGTCGC CCAGGCTGGA GTGCAATGGC ACAATCTCGG CTCACTGCAA
103701 CCTCCGCCTC TTGGGTTCAA GCAATTCTCC TGCCTCAGTC TCCCAAGTAG
103751 CTGGGATTAC AGGCATGCAC CACCATGGCT GGCTAATTTT TGTATTTTTA
103801 GTAGAGACAG AGTTTCCCCA TGTTGGTCAG GCTGGTCTTG AGCTCCTGAC
103851 CTCAAGCGAT CCACCCGCCT CGGCCTTCCA AAGCACTGGG ATTACAGGTG
103901 TGAGGCACCG CGCTGGCCAA ATGATGGTGT TTTGATCTGG GTCTTAAAGG

FIGURE 3, page 27 of 57

103951 CAGAAGGAAG GGGGGTAGTA AATTAACTGT GCTGGGGAAG AGAGGGAGGC
104001 CTGAGAGTGA GGAAAGAATG AGGGGTGATT CCAGGTTTAG GAAAACTGGG
104051 CAATTTGTTA GATGATGGTG CCATTGACAG AAATGGGAAA GAACAAGTTT
104101 GGAAAGAAAA CTCAAGATCT GGCTGGTGAC TTGTATTAAA CTTAAAGCCT
104151 CATTTGTGAC TTGAGCAGAA GTAAGGACTT TCTCCAGTGT TCAAGAGCTG
104201 GAAGGGATTT TTCTAGCCTC CAGGCAAGGT AATACCATAA GTCCCAACAG
104251 TGATGCCCTC CCTGGGAATG ATCTCAATGG GAGAATCCTA TACCCTGCCT
104301 CCTCCATTCA TTCCTTGCTC TGATGGTGGT TCTGGCTGGC TAACCTAAGT
104351 TACTCTTGCC ACTAGTTAAC GCCTGTCCTT ATTTCTCTTG TCCCCACCTA
104401 AGATGTCAAT CAAAACAGCA CGAGCCATGC TATGTCACAT GACATGTTGT
104451 CTGTCCAGCC CAGAGCTTGT TGCTGATGGG GGCACAGACT AGATTTTGAG
104501 AGAAATCTCT CTGTTACCAC CCTTAACATT CCAACCCCCT CTAATAGCCC
104551 ATTTAGGATT TATCATACTG TTTCATCCAA ACCTTTCATG ACCTGATTTC
104601 TATTTCCAGC TTCAACCACC CCTTGGGTCA CCACCTGTAC TTATTGAGTT
104651 TCCCTAGTTT TCTGAATTAA TGACTGAAGA TGATAAGCTT CCCTTACATA
104701 TGACTCTCAA ACCACCAAAC TGGGATTGTT GTTACTCTTA GTGATAATGG
104751 TTGCTATTTA TGAAACTTTT AATAGGGAAC ACAAACCCTG CCCAGAAATT
104801 CATATAAATT ATTTCATTTA AGAACATCAC AAAGTAGGTG CTATTATTTG
104851 ACCTTACACG TGAGACTTGA AGAACTTTAG AGCATTGCCC AAGGTCACCC
104901 AGCTAGTGAG GGTGGAGGC GGGATTTGAA TCCAGCTCAT CTGTCTCCAT
104951 TACCTGGAAG AAGGAAGGCC AGAGCATCAT GGCCTTTCAC AAGTTGAAGA
105001 GCCACGGGCT TTCTACGGTA GCCAGCCACG CTTTTCCATG ACTGGGGTGG
105051 GTGTGGCAAG TGATGAGGGT TTGGAGTTCA TGTGGTGGGG TGGCAGGGAC
105101 CAGGTGTCTT GGTAACTGCT GTTGCATTCA CTTCAGGAGC AAAGGACCAG
105151 ATCTGATTCT GCAGGATCAA CAATATGGAC ACTGCAGGCT CTGTAGACAT
105201 CCAAAGCTCT AATGGTGACT TGGGGAAGCT CAGGAGGGCA GGGAGGTTGT
105251 ACCCATTTAG AATGTAAAGA TTCCTATTTT ATAAAAAAGA AAAAAAGGAG
105301 ACTGAAGGCC TCAGTCTCCT CCAACAAAGC CAGGCTGTGG GGTAGCAGAG
105351 TCTCAAAGGG TGCAGGCCCA TGGCCACTGC CCAGGGCTCC TGCTCAGGCC
105401 TCCTCACTCC CACAACTGAG GGGAGACCCA GTTCCACACC CACCCACCTA
105451 GCAGTGTCTC ACACCCACCG GGAGAGGTCT AAACATCTTC CCTGGGAAAT
105501 GGTCCCAAAA TGTCCCTGCA GTAAGCAACC ATCTGGAGAG GCCCAGGTCT
105551 ACATCTGTTT TTAAAGCTCC AATAAATAAA TAAATGAAGG AAGAAAAAAA
105601 GAAGAAGAAA TGCAGAACAG GGTGACTAAA ATTGGCATGT ATTTTTAAAT
105651 GTTTATATTA ACAAACTAAC ACCTTTTAAC ATGAAAAGCA ATATAATTGT
105701 GCTAGCCACA AAATCATCGT AGGACTGAGA AAGGAATCGT GATTCTGAGA
105751 GCCCTAGAGT TAATGTGATC CAGCTGGCTC ATCCCTGTGA CTGCAGAAGC
105801 CTGTTTGGAG ATAGTGTCAG TAGCTTTTCA GGCCCTCTGT GAATTGCCAG
105851 AATGTGTGAC ATGAGCCAAA TTTCCCCCCA GCATCCCCGC CGCCGCCACC
105901 ACCACCCCCG ACCCAACCCT CCCGCCGGCT CCCATAGAAT AGTCACTGCC
105951 ATACAGAAAA AGAGAAGTTC TACTATTTCT GGGCAAGATT TCCACAAACC
106001 AGTTTGTCCC TTTCTGCTTT CATGAAATAA ACCATTTGGA TCAACGTCAG
106051 CTGATTGCAA AAATTTTCCC TTGTCTCAAA AGCAAGACTG ATAAGGAAGC
106101 AAACATGGGA GGACCTTAGT GGCCGAGCCT TTATGTGTAT GTTATTTCAT
106151 TGCTCTCATA ACTGCCCTGG GATGCTGTAA GCATGATTCA TCCTGTTTGT
106201 TTATCAGTTA AATTATGTAT CCAAGATTAC ACAGCCTATC CAGGATTAGA
106251 ACTCAGAGCC CTCGGCTGTG AAGCTTGAGC TCTTTCTTTT CAGTCTTCAA
106301 ATATGATCAT GCCATGAAGC AGCACAAAGC CCAGGAGGAG CCCAGTGAGG
106351 CTGGAGGGGT CCACTGGCAG CCACTCTCCT CCGTGCCCCT GTGGTGTTGG
106401 GGCAAACTTG GATCTTTCTG AATCTTTTAA CTGTTTCCTT CTCTTCCCGT
106451 TTTTGTCTGC TGGCTGACTT GTCCTACACT CTACTCCTTG CTTATGATAC
106501 TTATTTTTCC ATCCACAGCA AAACAATTCA CATCAAGGTA ATTGATGATG
106551 AGGCATATGA GAAAAACAAG AATTACTTCA TTGAGATGAT GGGCCCCCGC
106601 ATGGTGGATA TGAGTTTTCA GAAAGGTGTA GTACCCTGTC CTCCACACTA
106651 ACACTAACAT TCTTCTCTCC TCTTCTGTTT CTTCCTCTCC AACCCATTTG
106701 TCTCCTCCTC CTCTTGTCTT CCACCTCTCT GGTTCCCTTT CCCTTGTCTC
106751 CTCTCTTGCT CTCTCTCCTG CTCTCTTTTC ACTCCTCCCT CTCCTCTGTC
106801 CTCTCTCTGC CCCCAGCTCT GTCCTAACAC CTGCCAGCCT GACACATGGC
106851 ATCCATACGA GGGATGCTCA AGACCGATGG TAATTGTTCT GGGATAAGGA
106901 AATGAGTATG GGGAAAGAAA GAGCCAAAAT GCTGGAGTAT CATGTGCGGC
106951 TCTTGGCTTC TCCAGAATGG CTGGGCATAA AGGGGGGAAA AGGGACCACA
107001 TAGCCCAGCA CCAGACAGAA GAGCAGCACT GAGAAACAGG CTTTCAGCAC
107051 AAATTTCCAT GGGGCAGTTA TTCTCAGGGC TAAACTTAGA GTCCCAGGAA
107101 GTTGAGAATC AATGTATTTG GATTACAGTT CATTCCCCTC CCAAAAGCAG
107151 GCTTTAGGAG CCACCTTATC TGCCATGTTG CTACTATCAA GACTTGTTTC
107201 TCCTCCTGAC CTTGAGGAAG CTGAAAGTAC AGGTTTGAGT TCCAGATCTA
107251 GGTCAAATAT CCATTTGTCT TCCTATGTTT TTCCTATTAA GAACACCCAG
107301 GTGTGGAGGC AGAGAGTTAG AATAGTGGTG GAGATCATCC TGACCCAAAT
107351 GGAAGCTTCC CCAAGAGGTC CATGGGGCTT CTCAGAGTGG ATGGAATCTT
107401 TGCCTTCAAC TTCAATGACC CCATACATCC CATGGCCTCC AATAGACAAG
107451 TCAAGAAGTC CTTTCCTGAA TAGATCATAC TGTGGAGCAG GGAGCTGCCA
107501 GTACTGAGGG CAATGTTCCT TCCCCTTCCA AGCTGTCCCT CATGCCCTCC
107551 AGTACATGCC TGTTGTCACA GAGCACCCCA ATCCCATCCC ACAGCAGAGT
107601 TCCTGCAGCA GAGAAACAGG CTCACACCTT GTAGACAGCC CTGGGGTCCC
107651 ATATCTAGGG CCAACAGAAA TATTCCCAAA AAAATGCCTC TTGACAATCA
107701 ATGAGCTTTC TCTTTTGTCC GCTGAGCAAG GTATAAAAAG ATGTCAAAAG
107751 AAGTACCCAA AAAGGTAATA AAAATGTACA GTCGTGCATC ACTTAGCAAT

FIGURE 3, page 28 of 57

107801 AAGGATACAT TCTGAGGAAG GTGTCCTTAA GCAATTTTGT CATCGTGGGA
107851 AAATTATAGA GTGTACTTTC ACAAACCTAG ATGGTGTAGC CTACAACACA
107901 CCTGGACTAT GTGGGCCTAT TGCTCCTAGG CTACAAACCT GTACAGCATG
107951 TGCTTGTACT GAATATTGCA GGCAACTGTA GCACAATGGT ATTTGTGTAT
108001 CTAAACACAT CTAGACATAG AAAAGGCACA GTAAAAATAT CGTAGTATAT
108051 AGCCTTATGG GACCACTATT GTAGATGTGG TCTGTCATTG AGCAAAACGT
108101 TTTTATGTAG CATGTGACTG TACTTGTAAA GTACACACAC CACAAATGCA
108151 CAGCAAGTCC TGTGCCCTAC AAGCCCCTTT GGGTCAGTCT ACTACATTAT
108201 AAATGGCAAA GCCGAGCACG CCCACAGAAG GTAGCAGGAA CATCAGAGGA
108251 TCTGAAGAGA CATTTAGGTA AATGCTCTTT ACCCTTTAGA GCATTTAGTT
108301 CTTAGGCCTC CCCTCCCCCA ATCTCCCCCC CGCCCCCCGC CAAAAAGAAA
108351 AAGAAAAAGA AAGCAGAAAA TTACAATTCT GGCTCACTAG TAGGACCTGC
108401 TAGCCACCAT TGTGATTCCA TGAAGGACCA GAAGAAACCA TATAGGAAGA
108451 ATCAGGCCCA CACGGCAACC TCTCCACATG ACAAAGAGCC AGTCTTTGGA
108501 GGGCAGTGAA TTTCAAGGAA AGTTTTCTTC CCTGGGTGAC TTGTTTTTAA
108551 AAGATGTTAT GTTTTGTTGA GATACCCAGA GATGAACAGA AACTTCCATC
108601 ACCTTGTGCC CCAGACCCAT GATAATTCAC ATTGAGGAAA CCAGTTTTGG
108651 AACACATCAC CCCTAAGTGA TAGAAGCCCA AAGGTGATTT AGAATTTGAT
108701 GATTTACATC ATTTTCTTCA CATTTTCCCA GAAATGCATC AGCTGTAAAT
108751 AGTAAAGGAT TCCTATGTAA TATTGTGGTT AATACATATT TATTTTAGTT
108801 CCCACCACTG AAGCCCTATG AGATAAAGAA TGAGAAAGAT CACACAATTC
108851 TACCTCCCTT TCTTCTCTCT CTCTCTCTCT TTCTCTTTCT CTCTCACTCT
108901 CTCTCTCTCT CTCTCTTCTC TTCCTCTGTC TGGTTTTCCT TCCTCATAAA
108951 TACTTTTCTT TTAAAATTTT CTTTCTGAAA CTCACAATGG AAGTGAGTAT
109001 AGACATAAAG AAGGGACACA AGCCCTGGGT TCTGTTGACA TATTCCCTGT
109051 TGTGGGAAGA CCCTGGGTTA TTCCCAGTGG GTTAGTAGTT TACCTGTTGC
109101 CCAGAGAAAT GCCACTGTTA TCATGTGACA CCCAGTGGAA TGTGCTGCCT
109151 GACTCACTTC CTACTAACTG TTGGCAAGGT CTAAAATGAC TCCTCCTCAC
109201 CATTACCCGC CTTCTGCCTT CTCCTCCCCT TCTGTCCTTC TGGCTCCCTT
109251 CCTTTGCCCA CCTTTCCTTG CCTCCTGGCT CCCTGCCCCC TCACCCGTAA
109301 GAACAACTAT GACCAAGAAG ACAAGAAAAA CTAAGACCAT TTATTACCTG
109351 AGAACAACAC AATCCACCAT GGTCCTGTTG AAAGCCACCA TGGTGGGACT
109401 GGACTGCATG TGCCAGGAAT GACGGGGAAT GATTTTAAAG GCTGTGCTCC
109451 AGGTGACCAA CCAATCTACC GACCCAGTCG ACACACTCTC TCTCTTGTTG
109501 TCCCTACAGG AAAACCATAA GGGTTAAAAT AGTAGATGAG GAGGAATACG
109551 AAAGGCAAGA GAATTTCTTC ATTGCCCTTG GTGAACCGAA ATGGATGGAA
109601 CGTGGAATAT CAGGTGTGAG ATTCTTTAAA AACAAAACAA CAAAAAAAAA
109651 GAAAGAAAAA TTAAAACAAA CTGAAAAACA ACAACAAAAA AGAAAAAGCA
109701 GCTATATTTT TGTCTCCCTC CTTTTCTTCC CTTCTCCTCC TTTCTCTTTT
109751 TGACCAATGG ATTTTTTTAT TCTTTTCCCT CCTGTATTCT CGCTCTCACC
109801 CTGTTTCGGT ATCATCTCTG CCTTCTTAGC CTTAGCTTAT TCCAAATTCC
109851 TCCTTTACCG CCTTCTGGGC AGCACTGCAG CCTCAACTCC TCATTACCCT
109901 AATGAGTTAT TTCCCTGTTT TGCTACAATT TTCAATTATT CAATTGCCAT
109951 GGGCCCCTGC ACTCTCCCCC ACCCCACCCC TACACTGTAA CCTGTAAATG
110001 TGAAAATTCC TTGGTGGGTG GGGAGGAGAA GAAAAAAAAG GAATGTGATG
110051 CGATGCATGC CTGTGCCCCT TCCTGCCTTC CTCCCCTGCC ACCCCTCACT
110101 CTTTAGCCTG GATTGAATGT GGGGGGGTCT GGGATGGGGG TTGGGGCCTG
110151 GGTTGCAATG ATGCTTTGAC AGTTTTCTGC TGCATTCCCC AACTTCCTTT
110201 GAACGCTTGG CAGGTTATTC ACTTGTGGAG TGGCCCATAG GCCCCTCTGC
110251 CCTTCGAGGA GGTAAGTGTA TTTTCTGGCT GTTTCACAGT TGGGCAGACC
110301 GTGGCATGGG AAAGTGTACC AATTGTCAGA AGCCACGGCT TCTGAGAGCT
110351 CTGAGAGAGA GAGTTGACTT CTGGGGTAAT CATGCAATCT GGAATTCTGA
110401 GCTATTCTTC CTCCTCTGGG CATCCCACCC CATGCCATTC TATGTTCCTA
110451 GCCCAAGGTT GGGTGCCTCA TTCAGGCTAC TTTGGGACAA TGCAACCTCT
110501 AAAGCAGAAA ATTGAGAGTT CCTGAAGGGA AGGAAATAGT TCCAGGTATG
110551 AAAATTCCCG TAGCCAGGGG CCCCAGAAAA GGACTGACAT TGGGCAGGCC
110601 TGGAGTGTTG ACTTGTGGAT TTTCCAACAG AAGAGACTCT AAATGATGCA
110651 GTTGGTGCTG ATCCCTGACA GACAGGTGTT GGAAAGGTCA CAGATGTCTG
110701 CCTTTGCTTG GCATCTGCAA GAGAAAGTAC CGCCCAGATC CCAAGATAGC
110751 CCTCATCCCA CACTAGAGAA GTGGCCTCAT CTCCTGCTTT CCTCAGGACC
110801 TGCATCTGAG AATACCTGCC AGGGGCTCAT CCCTAAAGGA CTGATTATGT
110851 TGCAACCAGG GTAGAAGTAA GGAAGGATTT CTTCCCTTGA AGAAAATGAT
110901 TGGAAGCCAC TACTTTGAAT GGCTTCCAAT CATTTGGAGG CATAGATGTG
110951 GGAATGGGTT AGGGTGCTCC TGGGAAATAA CAAGAGGACG TTCACACTCC
111001 CATTCAGGAG AGATATGCTG CTGGGAGCCT CCTAGCAAAT GAAGCAGTGA
111051 AATCCACCTG TTTGTCAAAA AGGGGTGATC ATACTGCAAT TAGTTCATAT
111101 TCATGTGACA AAGAGCAGCA TAAAACTTTC CACACGAGGA CAGAGCTAAG
111151 AGATTCAGCA ACAACATTCC CAAAGGATTC TCTACAGGCC TTCTCAGTGT
111201 GATTGGTCAT TTCTCATTGT CTGCTGGGGA CTCTCCTGCA GAGCTGACCA
111251 CTTCTGTGCC TGCGCTGGTT TGGACACACC TGATGCTCTA GGGGCAGAAC
111301 TCCTCTCCTT CTTCACTGCT GGTTCTCTTC GTCACCACTC AATAAAACGT
111351 TGCCCTCAGC CTGACTGCCA AAAAGTGCTG GAAGAAAGAA ATTATCTCTG
111401 GTTCTATTGT TTCCCACATT GTATTCTTGC CCAACTTCCA GTTCTTGCCA
111451 CCAACAATAT TCTCAGAGGT TGCCTCAGCA CCTGCCCTAC CTCATTCCCA
111501 CCTCCCTTGA GCATTTATTC CATGTATTCA TAATTGGTTG GAAGCAGCAG
111551 ATACCCAAGG CCAATTGTAA GTCACCTTCA TCAGTTTCCA CAGTCCAAGC
111601 TACTTAGATG CAAACGAAAG CAGCACATGT ACAGCGTACA GGAAGGAAGG

111651 CAGTGGTTCC AGACAAGAGG AAGAGATTGG AAGTCCATAC ATGCCTTTAT
111701 TCCACCAGTA AAAAGGCTCT TCTCTTATGC CTCCCTTAAA ACCTCTACCA
111751 ACAGCAGGAC AGAGAGTGAC CCAAGATAAG TCTTCAAGAG ACCTAACCAA
111801 ATGCAAATGT CTTTGGCTAA TCCCCATTTA AGGACATCTT CCTGTTTTGC
111851 ACAGATTCTT TGCCCAAGGA AATGTCAGCA ATGCCCTCGT GGAGGGAGTA
111901 GGTGAGAAGA CAAGGATTTC AGCAAGCTAT CTGTGTGGTG TGCCCCCAGA
111951 TCTCCCCAGT GACCGAGATG CCAAGATGAA GAGTGCCAAG AAGAAATTGG
112001 TCAATTTTCC AGCTGCCTAT TTTATTGTCT ATGTTTTCTA GGCGGTTAAT
112051 TTCCAGTTTC TTCAGTACTT CCCGTATTTT GACATTAGAC CATAAGGTGA
112101 AAGGTCATAA AACCTGATTG TCTAGACTCA GAAGCAAATG GAAACCCATC
112151 CAAATTTCCA GAATTCCCTG CTGTTCTCAG AGTGAGAAAC AGAACAGTGG
112201 AAATTGCTTT TCATTATCAC TACTGCATGG GAGAGTCTGA AACATTCAGA
112251 ATGGCATAGT CTTTGCATGG TCAAAATGAC AATTGCATTA AAAAAATGAG
112301 AGACTGGATT TGAAATAGGA GACTCTATTT TTGGCAAACA AAACAGACTT
112351 CAGAGTTGAG ATTAAAAGCT CTGGATGAGC TGGGGGATGG AAAAAAGGGA
112401 AGGAAAAAAG GGAGACTGAA TAGGAAACAC AGTTGCTCTG GAGTCTAGAA
112451 GTGGACTTCC GAGAGCAACA CTGAGCAACA TAATCAAGAC TGTTGGGCCT
112501 GGGCCTGGAC ATTGGAAGCC TTCGGATAGA AAGGAAAGCT CTCTGTCTCT
112551 CTCTCTCTCT CTGAAGAATG GGGCCTGTTT GGTCCTCCTT TTTCGACAAC
112601 CGTGGGCTCA TCTTGACAAG CTGCCCAGAT GCTTCCTAAT TACTCACAGT
112651 CCTATGCTCT TTCCAGCTTG TCCCTGGGGT GTCTGAGCAG GAATAAATGA
112701 CTCTCACCTG ACCCAGGGGA TCAATACAGG GGAAAGTTCA GCTCCAGCTT
112751 CTCTCATGAG CAGCAGCAGG AAAAACACCC TCGAGGTATT GTGTCAGTCA
112801 AAGCTGGCCT ACCCAGGTCT TGCTGACCCA TCTATAACTG CTGAGCAGAA
112851 AGTCTTGGAT TCATGGAGAC AATGACCAGA GAATGATGGA ATTCCAGCCA
112901 ACTGCAGGCC TTCTCACTAC TCTAGGGATG GGCCAGATGT TCGGTGGCAT
112951 GTATGAGTGA AAACCAGGGC ATCAGGGACC TTTCTGGAAG AGCTGCCTTT
113001 GTCTGACCCA CCTGTGTTCA TTTATGTGCT GGGATCTCTG ATCTCCCCTG
113051 GAACTTGGGG GAAGCTCTTC CACGCAAACT CCCGGAAGGA GCAGAATAAA
113101 CAAGCTCTTG CCTATCTATC TATCTATCTA TCTATCTATC TATCTATCTA
113151 TCTATCTACC TATCTGCCTA TCTATATCTA TCTATCTCAA TGTAGTGAGG
113201 AAAGCCATTG ATCCATTAAC CTTTGGAATT CTACATGGGA GATACCTAAA
113251 AAAGTGAACT GCCTTGTTTA TGTATCATGC AGACTCTGGA TCCACATATA
113301 TCTCAGTGGC TGTGAATATA GGATGATTGA TCACAGGCCT GAGTTGCATT
113351 CCTACAGATT CTTAGGAAAA AAATTGATTC ACAGACATGT CCCCCCTGGT
113401 TCCCCCACAA CACACACTCC TTCCTCAGCA ATCTCTATCA GTCACCAACT
113451 ACACGTTGAA TATGTGGCAA GCTCTTCCCA GACCTTTATC TGAGAGCCAA
113501 GGAGTGAGGG GCTGTACTAA GATATCATAG AAATGAAAAT GTGGTGTGTC
113551 ACAAGTTTCC TTAATTCTTA GATCTTAAAC TCTAAGAGGG TTCAGCATAA
113601 GTACAAATTC AAGGGCTAGA GACAACCTGT ATTGGGTGTG TCTTTAACTC
113651 AGTTTCCCAA TCCACATAGG GACCTTGCAT TTGTCATCTC TCATCTATGT
113701 ATAGCTGTTG GTATGACAGT TTCTCTGTTC CAGAATACCT GAACTCTGAC
113751 TTAGCCTGTC CTTTCTGAAA CAGAAAAATC ACCCAACCAG AGATCTATGA
113801 GATCTATGGA AAAGACAGTT GCCAAAATAG ACAGCAAACA GCCAAACTTA
113851 ATTGAACACT ACCACATGCA GGGACTTTGC TAAGCAGAGG TGATACAAAA
113901 TGGGAGGAGC CCATAGCCCT AACTTCCAGG ATATATCTAC GGTAAAGACA
113951 AACCATTCAA GGAAAACATT CTGCAGGACT TACCTTTTTG CTAAGTCATT
114001 CTTTTAGGGG AAATCAAAGT TCTAGTCAAC GTGGCAGCTA GGAAGGCATT
114051 TGTGGTGATG GAAACCTTAT GAGCACTGAG AAGCTGAGCA TGAGTTCAGC
114101 TAAGTCGTTA GGGATGGAAG ACATAGACCT GGGCACTGTT CCACTCTTGC
114151 ACAATGCTAC CCATTTCCTT GAGCTCCCAT TCAAGCCCCA TGGTCATTTT
114201 TGCCACTCAT AAGTTAGCTA CTCTGGCAGG GTTGCAACTT ACACAGTTTT
114251 CATGATAACT GGATTCTCAC TCCTTTTTTT ACAGAATGGA TGTGATAACC
114301 TGGTATCCTA CACAGTCATG AGTGACCAAC CTACCCATTT GGTTCCCCAT
114351 CCTCATTCCT CCATTCCTAG CCCTAGGGTA GCCGGGAAAG CATAGGAGCA
114401 AATGCCCTTA CCAGGGCCCT GGTGCTCAGC AGCCTCTCCG GCTGCTCACA
114451 CCTCTTGCTG CTGCTCTGTG CATGCTCCAA AGGCTGCTTT TTGCGTATGG
114501 CTGCTGAGCT CTCACCTACT AAGCTCTCTG CTTTCCTTAT GCTGCCAGCA
114551 ACCACAAAAC CTGGTGATAC TTTCAAGATG GGACATTAAT GCTCTTTCCT
114601 TTTCTTTCTT CCATTTTTCT GGTATCCATT TGCAAACAGC GCTCCTGTTA
114651 TCTCCAGGTA AGAGGTGTCT TGTCCCCCTC TTTTCTTTCC ACTTCTTGCC
114701 AGTGCCATTA TTTGGTTTAA GACCAATGTC CTTTGATTTA TTGAATAAGA
114751 ACTGCAGGCT CAAGTTAACC TGACAATTTC TCCCAAGGAC TGGGAGATTT
114801 ATTTTCCCAC ATGAAGCAAT TATGAGAAAG CAATTGTGAG GAAGGCAATT
114851 CCTTGAGCAT CACTTCTGTC TGGGGACGTG GGTTAAGGCA TAGCTGATCC
114901 TCTCTGGGAC CAGGAAGAGA AATTAAGCTT AACAAGGAGA TGGTGGGTCA
114951 TAGACTTCTC CTGAGTCTTA ATTCATCTGC CATCTCATGT TGTGGGGGAA
115001 GAGACAGTGA GATTCAGAGC TGGAATCTCC TAATATAATT GTGACAGGAT
115051 TTGAAAAAAA AATACTTTAA TCCCAAGGGA TCCAGGAAAT AACCAAACCT
115101 GTTGTGAGAA TAGGAAATGC AATTTTTAAA GAATCTGGAA TTTTACCAGT
115151 CCTGGAGATC TTCCATCTCA TCACAGCTGA GACTTAAATT GCTAGAATTT
115201 TGGTTCATTT GTCATTGACC CTTAAAGTCC TATGTGCCGT GAACAAGATG
115251 AATTAGGATG GGGGATTGGG GCAGTGTTCT GGCTGGAAAT ATAAATTTTA
115301 GAGAATTTAT TTTGAAGAGA TTCTCATGCA GAATCTAGGT GCTATAGAGG
115351 ACGTACACCT ACTTTGAGAG TATGCTTGCA TGAGTGGAAA CCAATCATAA
115401 ACAACATTCA ACTTCATGAG CAGATATGAA AGCATTTTCA GCATATCTAG
115451 CAATACTATA ACTCTTTGTG CAAGCAGAGT GGCCTACACA AGACAGTTTC

FIGURE 3, page 30 of 57

115501 AATATATTTT AAAAGAACGT CTTACATTTC ATCAGTCCTT TGAACACAGA
115551 AAAAAATGTT AAGGCCACTT AAGAGGCAAA ACATCTTACA GAGTTCATTG
115601 ATATTCAAAG TCACCTACAG GCTACATCTT GGGTTCAGGA AGGGGCGGTG
115651 TACATAGTAA GGACATACGC CTTCTGGGAG CCTTAAACAA ACAAAAAAAA
115701 TGTAGGTAAC TCCTACATTT TTCTTTTGTG GAAAAAACAC AGTTACTCCA
115751 GCTTCCTTGG CTTTTTGCTT CTTTTTTATA CCAACAAAAT AAGGGCTATC
115801 CTCAACCCTC TGTTCTTCAT TCTTCTCCCA GGGTATTGAT TTCATAACAT
115851 TGGGTTTTTC TTCTCTACTT CACTCATCCT CTTGCCTGTG AAGGTATGTA
115901 AGGCTTCTTT GTTCCAACTC TTTCCTCCAC CCGCCCCCCC TCACATAAAT
115951 GCATAACAAA GATTGTGATT TAATTTAAGT TTCTTTCTAC TTTTAACATA
116001 TTTGCAAACA TCAATAGAAG CTAAAATGGG AAAAAGGAAA TGTTTCTTTT
116051 CCTAGCTCTT TCAATCTGTA AGCCTTTAAT TTAGGAGCGC TGATTAGCCT
116101 TTCAATTCGT TGGAAATCTC AAATACTGGT TTTAATTTTC CTAGGTGGAC
116151 AGAGACAGAG GGAATATGTT CATTCTGAGC TAACCACCCC CCCACCCCCA
116201 AGCTCAGGCG CCTTGCAGGA AGAGCACTAG CTACATCACT CTGCAGAGTG
116251 TTCACAACAT CCTATTCTTG TCTGGCCTGG CAAGCTCTTT GTCCTTCCAA
116301 TATTTGTTCA ATCTTCCATC CTATTCATAT TCTATCTTTC TCTCCCCTCC
116351 CAGCCTCTCT TCCTGTTCCT AGAACTGAGA GTTTATTTAG TCAGTCTGAA
116401 TATCTAGATC ACCTGCCATT TATTCTCTTT ACTTGAAATT CTGAGGAGTC
116451 ACATAAACAA GATATCAGAA TCACTATGGT CCTCTAAATT GAAGACTTAT
116501 AATTCTCTCA AGAAATTAAC AACATTTGAA TTTAAAGGAA AGATCATGAC
116551 AAAAATAGAA AAAGGCAGGA ATTATTGCCA AACCGAGAAA CTAGAAACTA
116601 GAATTAACTT AAAGGCATGT GACTCAATCA ATTAACAAAT ATATACAGAG
116651 AGCCTCTGTG GGACTGTGGG AGATCCAAAG ATAGAGGATT GGTTATTTGT
116701 CAAAGGGATT TTTGCAGAAA GCTAGATGGA AAAACTGACT GTCACCACAG
116751 AGGTGGACAG GTCAGTAAGT AGATCAATAT CCTGCCAGAT GGATATAGTG
116801 CTAGATTGAT AGGTAGACAA GGGGTTAGAC AGGTACATTT ATATGTCACT
116851 GGAGAGCTCA TTATATTGGT ATAAAGTTAT TGTGTCACAT GTAAAGTATG
116901 ACATGGGGGA ATTGGGGAGG AAGGAGTGGA ATAATACTGT CGCTGCTAAG
116951 ATAGGCATTG TGATATGGTG CTTAAACCTG CAAGTAAAGG AAAAGAGTAT
117001 GGAATCTGTG TGTCTTTTTC TAAGGGCTTT TTCCCAGAGT AGCTTGCAGT
117051 CTGGCTTCTA GGGTTGCTGG CCTATAGCCA GAACCCTAGA TTCACCCAGA
117101 TTTACCTTCA GAATTAACTA ATCAGAGACT CAAATTCAAT AGACTAAATG
117151 AAGTCAGGCT GCTAGAGGAT GTCTGCTGAC TTGGACATAT GCAGAAAGAC
117201 ATGGATCCTT GAGAAAACAT TGTTTCCAAA AGTGGCCACC AGCACTAGAG
117251 GAAGGACAGC ACCACGGACA GCTCCCAGAC ATTTTAGGAT TGCCTTCTGT
117301 GTTTGGTGCC CGAACACTGA GCAAAACAGC GAACTCAGGA AGTCTCCACA
117351 CACTCTCATA CCATCTTCAT GCAGTCCAAC TAAGAAAATT CTTACATAAA
117401 ATATAAGGCT GTCTGCTTGG TAATTTAAAC CCTTGGCTTA TAGTCTTTTC
117451 AGTGAATTTC TTTCCTTGCA AACTCGAGAG TTGGAGTCTC ACGACTGCCC
117501 TTGCTTCACC AATTCCCCAG CTAGAGACAA AAGACCTTCT TGGCCTCTGA
117551 CCCATTTTGT CCTTGAGATT ATCCAAGGAC TACAGGATTC CCCTAGGAGG
117601 TTTACTGTGT GGAATGAAAG CAATTAAGGA GCTGAATAAA AGAAATAATT
117651 GCATGTGAGA ATGTGGACTT GGATGGGAAG ATGTTTAAAT GAGCTCTGAA
117701 AGAAACAAGC TGCCAAGAGC AATTTTCTAA TTAAAGGGGA ATAAAAAGAT
117751 TCAATCTCTA TTTCACTCTA ATCCAGAAAA CATGTCTTCA TGGAGAAGTG
117801 CTCTTAAAAT GGACTCATCA GCCAAAGTGG AAAAACAAAA AACAAAAAAA
117851 CTGTTCAACA TGAGAAGGGA CCATTGGTAA ATGAGTCAAG ATGCTGTGAA
117901 ACCAGTAGAC ATTTCCTTTG AATAAATGTA CTTCTGCACC TTCAAGAACT
117951 CTTACAGGAA GTGGTTGAAC AAACAGGCCC AAAAGTTCAA AATAGTTCAA
118001 GGTCAAAACA CTTGCCCTTT CTTCCCAGTT CCCCAACATC TCACTGAGTG
118051 TCTTGAGAAC TTCACTTGAT GCTATTTCTC AGGAGATGTT TAGGTCAGGT
118101 TGTCCACCCA GGTATAAAAG AGAAAGAGGA ACGCTTATCC CAGTCTGCAA
118151 GGCACATTCT CATGGTCTGG TTATAAAGTG TTTAGTACTT CATAAAAAAG
118201 GCACTAAAAA TATATATAAA CTCCCCATTC CCAAGAGTTA TTTGCTTTGT
118251 ACCCACTGCC CATGCCTAAT ACTCTGAGCT GTATCCTTCC AGGGAATGGA
118301 AAAGGTGTTA AAGCGAGTCT GATTTTGTTT TGTTGCAGAT GTGACAGACA
118351 GGAAGCTGAC TATGGAAGAA GAGGAGGCCA AGAGGATAGC AGAGATGGGA
118401 AAGCCAGTAT TGGGTGAACA CCCCAAACTA GAAGTCATCA TTGAAGAGTC
118451 CTATGAGTTC AAGGTCAGGC AAACAGTGAG GTCTAATTGA ATAATAAATA
118501 AATTAAAGTG GGAGGCAGAA GACCTGGGGT GTTTTTTTCC ACTTTCACTA
118551 GTGAATATGT GAAGTTGAAA CTGAACAAAT CACTTACCCA CCCCAGGTCT
118601 CAGTTTCCCC ATTTGTAACA TGAAACAAAT AGTGCTGACC ATTTGTATGC
118651 TAGGAATATT GTTAGGAAAC ATAATATAGA ATGTGAAATA AGTGGACTAG
118701 AAAGTCCTGA GATGTATTAT CATTATTGTT TAACTGTGTT TTTAAAGCAA
118751 AAATATTAAA ACTCACTACT ACAGGGCAAG ATATATTAAC ATCATTATTA
118801 TTATTCATTA TTGTATTATT CTAAATAGCC AATTTCAAAA GTCACAACCA
118851 GGCCAGGCAG TGAGGGACTC ACGCCTGTAA TCTCAGCACT TTGAGAGGCC
118901 GAGATGGAAG GGTCACTTAT ACCTAGGAAT TTGAGACCAG CCTGGGCAAC
118951 ATAGGGAGAC TCCATCTCTA TAAAAAATAA AACAAAATAA AAATCAGCTC
119001 AGTGTGGTTG TACATGCCTG TGGTCCCAGC TACTCAGGAG GCTGAGGTGG
119051 GAGGATGGCT TGAGCCCAGG AGGTTGAGGT TGCAATGAGC CATGATTGCA
119101 CCACTGCACT CCAGCCTGGG TGACAAAGTG AGACCCTGTC TCAAACAAAA
119151 CAAAACAAAA AGATTACAAC CAAAAACAAA GGGAAATAGA AGGATTGCCT
119201 CAAAAGAGAT CGCCCAAGGC CATTCCATGC GTAACTGTCA GAACACCTTG
119251 GAGACAGGGC ATCTTTCATT CCTTTGAAGA ACCAGACTCC TCATTGGTTC
119301 TGAGCATTCT AACCTCATGG TTCCAAGTTT TTCTCTTCTT AACAGACTAC

119351 GGTGGACAAA CTGATCAAGA AGACAAACCT GGCCTTGGTT GTGGGGACCC
119401 ATTCCTGGAG GGACCAGTTC ATGGAGGCCA TCACCGTCAG TGCAGGTGAG
119451 AAGTGTCTCA GGCTGGCCTT GCTGGGAGAA GCAGGCAACC TCTGAGAAGG
119501 AAGCGTAAAG CCACGTTAAC AGCCTGCCAG TCCCTAGGAA GGCTTGTGTG
119551 TTCAGTCTTC CCAGCTCTGG TCCTAGGTGC CTGCTTGGAA AAGAATCATG
119601 GCGTATCTGA AAAACATGGT TATCTCTGGT TTCAAATCGT TGTTCTGCTG
119651 TGTGAACTGG AACAATGTAC CCTCTCTGAC CTCAATGTCC TCTTTCCAAA
119701 GGGGAACTAT TGCTACCTTT CTCAGAAAAG TAGAAAGGTA CAGAGTCTTG
119751 TATAAAATCC AAACTCAATA AATTCTGATT TCTGTCATTC TTTCTTTTCA
119801 TGGGTTTGGT CCCGCTCTTC TGTAAAATGT GGGACAATTC TGATTTAGAG
119851 ATGTGGGAGT TAGGAGTTTA TAAAATGTGT TGCATTGACT CTCCAACAAA
119901 ACACTCTGGA TGATTCCATA CCCCTCCCTC GGCATTTACT GACAGGCTCC
119951 CTCAGTAGTG ACCCACAGCA CAGCCGGGAG TCCTAGCAGC CTGAGGGGAC
120001 TGCTGGTTGG AACAGGGACG GAAAAGGTCT CCCAACCACC ATCACTATCA
120051 CCTCTCAGCA CCACTGAGGC CTCCTGGCCT TGTCTTTTAT TGAGAGACTT
120101 TGTTGTCATA GCAACCCACA GGGTCATATC CCCAAGGCCC CAGAGCCAGA
120151 GCAAAAAGAC AGCCAGGAAG AGAGGTTTGC TGCTGCTGCT GCTGCTGCTA
120201 CCCCACTTTT CTCATCACCT GCTTTAGATC TTTCTAGCTC CCCCTCTGAT
120251 GACCTGACTG TGCCCCTCAA GACAATAAAC GGAATGTAGG CCACATCATC
120301 TACCCTGCTC CTTTTACAAA GGAGGGGACT GAGGTTCAGA AATAAGAGAT
120351 GATTTACCCC AGCTTACAGA TTTTCTTCAT GGCAAAGCTG GAATGAGAAC
120401 CCAAGTGTTC TGACTCCTGT TCTTTCAAAA CCCAGCTTCT ACCGGTTATG
120451 CCAAAACATG ACAGAAGTTG CCGTTGGCAA GGCACAGGCA TGCCTCAGCA
120501 TACCCTCCCC TCCAGGGCTG CTGAGTGGGC AACTCTGCCC ACATTTCCTG
120551 GCAAGGACAA TCAAGGCCCA TCCTGCTTTT TCCCATGAGA TGTTTGGAGG
120601 AGGGCACTGG CTCTGCAGTA TATTCTCGTG ATCTGGAATG ACAGCCATCC
120651 CTCAGGGGAC AGATAATGAC CAGAACCACA ATGGTTATTG CAGCAGTCAG
120701 GTCAGAAAAT TTGAGAGGAG CCCTGCTGGC ATCCAGTGAA GAGTGGCCAC
120751 ACCGAACTGA TTTCACTTCT CTCCTTAGAC AACAAAATGC AGCCTGTGCA
120801 TTCTCCTTTC TTTTTTTTTT TAATTATACT TTAAGTTCTG GGGTACATGT
120851 GCAGAACATA GAGTTTTGTT ACATAGGTAT ACACGTGCCA TGGCGGTTTG
120901 CTGCACCCAT CAACCCGTCA TCTACATTAG GTATTTCTCC TAATGCTATC
120951 CCTCCCCTAT CCCTCACCCC TGACAGGCTC CAGTGTGTGA TGTTCCTCTC
121001 CCTGTGTCCA TGTGTTCTCA TTGTTCAACT CCCACTTATG AGTGAGAACA
121051 TGCAGTGTTT GGTTTTCTGT TCTTGTGTTA GTTTGCTGAG AATGATGGTT
121101 TGCATCCTCC TTTCTTTCTG CTCCACTGTC TTGTCCCTCT TAATCTCCTT
121151 CTTTCTTCTC TTCCTTATTC CCTGGCCCTC TCTCTCCCAC TCTACCTTGG
121201 TGCCCTGCAT TCAAATTGAC CTATGAGGCA GCCCAAATTG TTTCCCCACT
121251 ATTTTCTGGC ACGCTGGCCC TGGCCCCCAC CAGCTGCCCA GAAGACAGCT
121301 GGAGTCCCCT TCTAGCGGAT GATGCCTGTG GTGCGGGTTG GGCTTGACTT
121351 TCTCATGAAT GATTATCTGA CTTCTTACCC GTTCTCTTGC CTGTTTATCT
121401 TGCCTTCAGC AGGGGATGAG GATGAGGATG AATCCGGGGA GGAGAGGCTG
121451 CCCTCCTGCT TTGACTACGT CATGCACTTC CTGACTGTCT TCTGGAAGGT
121501 GCTGTTTGCC TGTGTGCCCC CCACAGAGTA CTGCCACGGC TGGGCCTGCT
121551 TCGCCGTCTC CATCCTCATC ATTGGCATGC TCACCGCCAT CATTGGGGAC
121601 CTGGCCTCGC ACTTCGGCTG CACCATTGGT CTCAAAGATT CAGTCACAGC
121651 TGTTGTTTTC GTGGCATTTG GCACCTCTGT CCCAGGTGAG AGTGAGAGGT
121701 GCTTGAATTT GCAAAGAGGA TTTTACCTGG TTCAAATGAC CCCTGGACTC
121751 CATCTCATTA TCTTCCACAC CATCTCAGAT CTGAACTTAA CAGAGCCTCT
121801 GCCCTTAAAG TGCACAAAAG TCAATCAAAG AGATGAATAA TGACATTAGT
121851 AATGACAGCT AATATTTCTT GAGCACTTTC AATGTGACAG ACACCATGTG
121901 TGTTCAGCAA TTTACACATT TACATTTTCC CCCTGTAATG TTTCCCAAAG
121951 CCCTATTAAA TAGGGTAAGT TATTATCCCC ACTTCACAGA CAAAGAAACT
122001 GAGGCCCACA GAGGTTAAGC TACATGCCCA AGTAAGTGGT CCAATTTCTT
122051 AACCTCCACA TTATGTGAGT AGACCACAAA CAGTGAAATT AAAAGAATGT
122101 AGATATTGTT CTCCTTCTAT TTACCTCTGG CGATCTCTGA GAGGTTAAAG
122151 ATTAGCCAGC TCAAAGATAT CAAAGGAGAA ATGCCCACAT ACATTCTTGG
122201 CCTCCTCTAC TTGGAAGGAC ACTGTGAGTA CAAAGTATCT CCTAGCAGGA
122251 CAGCCAAAGG AAGTTCCACA GCTTTTATCT TTTTATAGGA TGAATTACAT
122301 ACTCTTTCTT TTTCTTAGGA ACACTCAGAG ACAAACAGAA AGGAGCGGAC
122351 ATTCCTTTAC TCATTGAACA AATATTTACT GAGCACCTAT TATGCCTGTT
122401 ACAGTATTGT GCTAGTTTTT GGGACTATAG TGAAAGGCAA GATACACATG
122451 CTTCCTTCTC CACGTGGAGT TTATAATCTA CTGAAGGAGG CAACTCTCAA
122501 CTACTGTAAT TAAAGTTATC TTGTTAAATC CTAGGAAGAA AAAGAAAAGG
122551 TACTGCATAC GGAAGGAAGT TGGGCCTGAA TGTAGGAGTT AGCAGGTAGA
122601 CAGGGGCTGC ACTAGCCCAG GTTCTTTACT TAATTCAGTT AGGGGCTTTG
122651 GGGCCTCTGA ACTCTGAACT TCTGCCAGGG AGCTGGCATC CCAGTTGCCC
122701 CAGAAAGAAA CAGAGCACAT CCTCCTGCAG GGAAGTTAGG CTGAATCTCA
122751 TCAGACAGGA CTTTTCTGGC TGGGCCAAGG GAAATCTTTC CTGTACCAAG
122801 CAAACATATC CTTCAAGAGA GTAGCTGAAT TCACATCAAA TTCTAGGAAA
122851 ACCTCTTTCC AAAACCCCAG CGCAGGCCAG CGGTATTATT TGTCCATTAG
122901 TGATGCAAGA GATTTAGCTA TCGTGGAAAT GCATCAGAAG GTTGGAAATT
122951 AGATGGATGA TCCCAGGAAG GCCTGTGGAT GAGATGCCCT GTGATCTCTG
123001 TTCTCCAAGC CTTGGGGGAC CTGAACTATC AGAGGGGAGG GAGGAAATAT
123051 GGGGGAAAGC ATAGAGGTGG GAAGAAATAT CAGAGGATCA GAAGCAAAAA
123101 ACAACAATAA CAACAGAAAC AAAAACAAAC AAACAAACAA AAAAACAAGG
123151 CCATAGGCAA GAAAGGGTAA GAGGTTTTCT CTGGGAGATC TAAAAAAAAT

FIGURE 3, page 32 of 57

123201 GGCAATAATG AGGTAAGCCA GGCAGATACC TTTGGGCATC TCCAAGTCCT
123251 TGCAATTGGC CAAGACAACA GCTAACAACA TTTGAGGCTT TAAGAAGGTT
123301 ACCCTGTGAT CCACTCATCT GATTTAGTGG CTTTGGCTGA AGCTCTTTGG
123351 ATATAGTTGA AGGTACGGAA AGGGTCCTTA CATGAGGACT TTAGGGTCAA
123401 GTCTCTTGCT AACATCCTAT GTGACCTTGG GTAAATTCTT TGACCCTTAT
123451 TTTTCTTACC TGTAAAATAA AAGAATTGGG CTAGATGTCT CTGACAGTCC
123501 TCCCTGTATC TACAATCTGT GCCAAGATCT AAAGTCAAAC ACCCTGCAAG
123551 GCCCTGTGAT ACATATATAA ACCACAAAGA CAGAGCCCCG TCTTCCTTGA
123601 GTCCACAGTT CACCCTGCAT GTCCCCATCA TGGTTCCCCA ACATGTCCTC
123651 TGTCCCCAAA ATCCAGCACC TCACCCAGTG CTCAATCAGT AGGCATTGCT
123701 CAATAACTGT TGGTGGTTCG TGAATAAATG CCCCATATGA CAGTTAAAAT
123751 CAGGCATCTA CTCCAAGCAG CTTCCCAGGG TGTCAAGGTT CCCTGGGGAG
123801 ATATTATGGG ATGGCAAACT TCCCTTACTG AAAAAGTAGT CAAAGGAGAA
123851 CAATAAGCCC ACTCAGTAAA TATCAGAACT GGAAAGCCCT TCAGAATCTT
123901 TCAGATCACT GCAGATGAGG AATGGGAAGC CCAGACTAGG GATGTGACCT
123951 ACCCAGGGCC ACACGGCTTG CTTGCGGCAG AACTAGGAGT TAGGAGTGGC
124001 CCCCTAGCCC TTGTCTCTCA TTCCTGGGTT CAGCCCACCA GCTCAAGCTG
124051 CTTTTTGGGC ATACTGGAAG ACAAGCCCTG CACACCTTAG CCTCCTACCA
124101 GTTCCCATGT GTCTTTGTCC TTTTCCAGAT ACGTTTGCCA GCAAAGCTGC
124151 TGCCCTCCAG GATGTATATG CAGACGCCTC CATTGGCAAC GTGACGGGCA
124201 GCAACGCCGT CAATGTCTTC CTGGGCATCG GCCTGGCCTG GTCCGTGGCC
124251 GCCATCTACT GGGCTCTGCA GGGACAGGAG TTCCACGTGT CGGCCGGCAC
124301 ACTGGCCTTC TCCGTCACCC TCTTCACCAT CTTTGCATTT GTCTGCATCA
124351 GCGTGCTCTT GTACCGAAGG CGGCCGCACC TGGGAGGGGA GCTTGGTGGC
124401 CCCCGTGGCT GCAAGCTCGC CACAACATGG CTCTTTGTGA GCCTGTGGCT
124451 CCTCTACATA CTCTTTGCCA CACTAGAGGC CTATTGCTAC ATCAAGGGGT
124501 TCTAAGCCAC ACAACAGAGC CTCCAGCAGG GCAGGCCTAG GACTTCTCCT
124551 AAGAGAAGGG CACTTCCCCA CCAGTGATCT CTCCCGACTG CACTGCCCTG
124601 GAGAGGCAGC ATCAGGACCT AAGCCCCAGG AACTTCACCC AACTTAGGCC
124651 CTGGCAATTA ACTGAAAGGG CAAAGTCTTA ATCAATCAAA CAATGGAGGA
124701 ATCACCGACT TTACACAGTA TTTAATTGAA TACAAACAAG CAACAGCAAC
124751 AAATCCACCT CCACCCCATC TCCCCCTCAT ATCCCTGACC CAAAGCAAAG
124801 GTCAGAGCCT TTCGCCTCCT TCTATTCCAT CTTTTGATTA TTCCTTTGCC
124851 TCTCATTTCT TTGGAAGCAG GGTTTCTCCT GTCTGCCCAA TGCCATATGT
124901 CCCTATTATC TCACTCAGCT GACAAGACGT GAAAATGAGT CACATTCATG
124951 TGGCTGGGGT GGGGTTCTTT TTTCATTGTA ATCATTATTG TGGTTGCTTT
125001 CGTTTTGCCG TTAGGTTTTG CTTATTATTT TGTTTTGTCT TTTTTTTCTG
125051 AAGTGAGTGA AAAAGGTGCC ACAAAGGAAT TCCAGGTCCG AGCCAACAGA
125101 GAGAAACATG AATTTTTAGA CACATGCTCT CCTGCCACCT CTTGGCTCCA
125151 TCAAGATCCA GTTCCCCATC TCACTGTTTT CTCTGAGTTC TTGGGAGGAG
125201 TGATGGTGTT GGGGTAGAAA TAAGCTCACT CACCCACGCA GGGTACTAAA
125251 GATCTTACAG GAGCTTCAAC TGGAGCAGGA GGAGCTTTTT ATGCTTATGT
125301 TGAATCAAGT CAGATACAAA AAGCAATTGT CCCTCTTTGC CCAAGCCTTT
125351 CCAATTCTGT GTGTCTTGTT GTGTCAGTGT CCACTTGTGT ATCCTTCTGC
125401 AGGAAGACCC GCCAAATAGA AGAGATGGGA CAAAAATAGG AATGGTGTGT
125451 GACGACAAAG GGCTACTGGA AGAACAAAAG GGATACAGGC CTTCTTGATT
125501 ATCTTTGGCT TTGTACCTGA GGCAGGAGAG AAGAGATGTC CAACCAGTGA
125551 GATCTTTAAG AGAAAAGTTT GTATTTTAAA TGTCAATGTG CCTGAGAAAT
125601 GTCAGCTTCA CCACGCTCTT GCTTCCTAAT GCTCTATACA AAGAGGGCTG
125651 ACTATATTTC TTGAAGTGGT GTAAAAACTT AGAGATTTTA TAAGAGAACC
125701 AGGGGCTCCC TTCACCTCTC CTGGTCCCTC AGGTCACATA TGAAAGCATT
125751 TTTACAAGAT AGGAACTGGA ATTCCTCATT TCTCCCATGT TCCTGCTTGT
125801 TCTTAAACTT CATGAAGCTA TTTTTCCAGC CTATGGGGTA GTTCTTGCTC
125851 CAGTAAGAGG AATCTTAGTT GTCATAATCC CTTGGAGCCT GGGTTTTTGG
125901 AGAAAGAGAT CTCCGTGCCC TACAGACCTT TTCTCAACGA ATGTGGGAAG
125951 GACCTGGCTT TAAAACACGC ACACAAACAC ACAAATAAAC AGACATAAGA
126001 TGTCATCACG AAACTGCCCA CGGATCTTTA GGCTTTCTGC ATTGACATAA
126051 ATACATTTTC TAAGGGGGGG GGGGAAGAAA TTAAAAAACA CCTGTTAATT
126101 TTAAACACAT TTTTTAAGAA AAAAATAATT AAAAAAGAAA CAGTGCTCAT
126151 GTCATAAGCT ATGTTGACAG TTGCCAGTGG AAATGTTGGG TTGGTTCAAA
126201 AAAAAAATAA AAGCTATACT ATATCTCTCT ACATACAGCT TGCTTCTACC
126251 TGTGTTTCTT CAGTGAAAGG TCCAGGGGGC CACTGTGGGC TTCTTGTGAG
126301 GAGACGTGAC TCAGGTGAAG GTGTCACCTC CTCTCACACT CAGGTGCCAA
126351 TGTGTCAGAC CCAGTATATT CTAAGCAAAA ATACTTCAGG AAAATGCCAC
126401 TTGTCAAAAC CTGGACTTTG CGAAGTTGGA AGATGTAAGT AGTAGTAAAA
126451 GCTGTGGTAA TTATGGAGGA AGGAGGTTTC TGTATCAGAA AGGCATTGGC
126501 CGTGACAGAC TC
(SEQ ID NO:3)

FEATURES:
Start: 2010
Exon: 2010-3793
Intron: 3794-109509
Exon: 109510-109613
Intron: 109614-118338
Exon: 118339-118463
Intron: 118464-119345

Exon: 119346-119445
Intron: 119446-121409
Exon: 121410-121685
Intron: 121686-124128
Exon: 124129-124502
Stop: 124503

SNPs:

DNA Position	Major	Minor	Domain	Protein Position	Major	Minor
378	C	T	Beyond ORF(5')			
742	T	-	Beyond ORF(5')			
2005	C	T	Beyond ORF(5')			
2381	A	C	Exon	124	T	T
5165	C	T	Intron			
5402	A	G	Intron			
6794	T	C	Intron			
9883	A	G	Intron			
10210	T	C	Intron			
12220	T	G	Intron			
13842	G	A	Intron			
14200	C	A	Intron			
15878	G	T	Intron			
16030	A	G	Intron			
16292	T	C	Intron			
16506	T	G	Intron			
17953	C	A	Intron			
23832	C	G	Intron			
25001	C	A	Intron			
25141	A	G	Intron			
25191	A	G	Intron			
26147	-	A G	Intron			
27400	A	G	Intron			
27401	A	T	Intron			
29278	C	T	Intron			
31437	A	G	Intron			
31857	A	G	Intron			
33155	G	A	Intron			
39487	G	C	Intron			
41449	T	C	Intron			
42420	T	C	Intron			
43256	G	C	Intron			
43967	T	C	Intron			
48604	-	A	Intron			
49560	A	T	Intron			
52729	G	T	Intron			
55031	A	G	Intron			
55066	A	C	Intron			
56912	A	G	Intron			
58480	C	T	Intron			
61128	G	A	Intron			
61320	G	A	Intron			
61444	A	C	Intron			
62641	T	C	Intron			
63023	A	G	Intron			
63051	T	C	Intron			
64989	T	G	Intron			
65929	C	A	Intron			
66694	C	G	Intron			
66755	T	A	Intron			
66879	T	C	Intron			
69156	C	T	Intron			
69280	C	T	Intron			
70647	C	T	Intron			
71867	C	T	Intron			
71900	C	T	Intron			
71901	G	A	Intron			
72369	C	T	Intron			
72992	T	G	Intron			
73154	-	T	Intron			
73164	-	T	Intron			
74149	T	A	Intron			
74171	G	A	Intron			
74918	A	G	Intron			

75386	G	A	Intron
77751	G	A	Intron
78264	G	T	Intron
80986	T	A	Intron
83609	C	T	Intron
85271	G	T	Intron
87770	C	T	Intron
87837	T	C	Intron
87866	C	T	Intron
88238	A	C	Intron
89219	A	G	Intron
89331	T	C	Intron
90794	A	G	Intron
92404	C	T	Intron
92672	A	C	Intron
92684	A	G	Intron
93132	G	C	Intron
93537	A	T	Intron
93557	T	C	Intron
95067	C	T	Intron
96000	T	C	Intron
96877	G	T	Intron
97271	A	C	Intron
97470	G	T	Intron
97518	G	A	Intron
98476	C	T	Intron
98779	C	T	Intron
99218	C	G	Intron
100538	C	A	Intron
101045	A	C	Intron
101232	C	G	Intron
101266	G	A	Intron
101290	A	G	Intron
101326	G	A	Intron
102342	C	A	Intron
104489	C	T	Intron
105266	A	G	Intron
105338	T	C	Intron
105570	C	A	Intron
105928	G	A	Intron
106459	G	C	Intron
107710	C	G	Intron
108062	G	A	Intron
108214	G	A	Intron
108364	C	A	Intron
108657	T	A	Intron
109746	C	T	Intron
111484	G	T	Intron
112879	A	G	Intron
113245	C	T	Intron
113265	T	C	Intron
113497	C	G	Intron
114486	G	T	Intron
114686	T	C	Intron
114817	C	A	Intron
115600	G	T	Intron
115668	A	C	Intron
115745	A	G	Intron
117230	A	C	Intron
118908	A	G	Intron
120430	C	A	Intron
120830	A	T	Intron
121926	T	C	Intron
122102	G	C	Intron
122950	T	C	Intron
123366	C	T	Intron
124947	C	T	Beyond ORF(3')
125010	A	G	Beyond ORF(3')
126043	T	C	Beyond ORF(3')
126064	-	G	Beyond ORF(3')
126283	C	G	Beyond ORF(3')

Context:

DNA
Position

378 TGGCATGTACAAAGGTCCTGGGGTGGACAGTCACTTGGTATAATCCAAGAGTGAACCTGA
AGGCTATTGTTGTTGAAATGTAATAAGGGAGAGAGTGACGGGATGAAGGGGGATGAGTGG
GAAGCAGTGAATTCCTGCAAGGCTTTGAAGGTCATGGGAAAGAATTTGGTCTTTATATCA
AGAGCAAGAGAAGACTACTAAAGGGCTTCAAACAGGGGAGCGATATGCTTAAGTCTGTTT
GTTTGTTTTTTTAAAAAAAGATTACGGTGGCTATATGAGGAAAGTGGAATTGAGAACTAG
[C,T]
GAGAGTTGGAGTGGTGAGCTCCATTAGGAGGCTACTGAAGTAGATTCATGAGGTAAGGAG
TGATGGTGGCCTGGGCTGGGATGATGGTGGTAGAAATGGAGAAAGAGTTGATAGGATTTA
GTGATTGGATAAGGGACAGAAGAGAGATGAAGGCTTTCAGACTAACATCTGCTTTCTAAC
ATGAGTAACTGGGTGGCTGAAGATGCTATTTTCTGAGCTGGGAAACAGGAGAAAAAGGAG
CAAATATGGGGGATGAAGACTTTGAGTCTTTAAGGTGCTGTACAAACACAAATCAGCATT

742 TGGTGGCCTGGGCTGGGATGATGGTGGTAGAAATGGAGAAAGAGTTGATAGGATTTAGTG
ATTGGATAAGGGACAGAAGAGAGATGAAGGCTTTCAGACTAACATCTGCTTTCTAACATG
AGTAACTGGGTGGCTGAAGATGCTATTTTCTGAGCTGGGAAACAGGAGAAAAAGGAGCAA
ATATGGGGGATGAAGACTTTGAGTCTTTAAGGTGCTGTACAAACACAAATCAGCATTCCT
TTATTACTAAGGGTATCCCACACAGTTGTAGCAGAGGGAGAAAGATCGCCCCCCCCCCAC
[T,-]
TTTTTTTTTTTTTTAGCTATTCCATGGTATTTTCATTCTCATCCCACCCAAATGAGGCAG
TGAGTGGTAAGATGAGTATATAATAGTTTCAATTGCATTTCATCCCATTCTTCTGAGCTC
AAGCTCACCTTTTAGTGGTTTGAGGCCAGTAGATGAAGCTGCATATCACCCCCAAAATCT
TGTCTCTAGTTTAACAAAACTTATTTGAGAGACATTTGCATGTTTTATTAATAATGATTT
TTACCACTTGTTCCTTTCCATGTTTGGGTTTGAAATTTGAGTGGCTGGCGGATGATCATC

2005 TTTCCATCCCCAGTATTCCCAGCTATTTCAAGCCATTTTTCAACGGAGTCTCCACCAGAT
GGTTTGGAGGACAGAGCAGCTATTTGTGCCTCCCATTGACATCTATTTTTCCAAGTGAGA
GACTGCCCCATATGTTAGTGCAATATGTCACTGGAGGTGAAGCATCAGTTGTATTGGTGG
GAACCTGCCGTTTGCTGTCCCCTTTTTCCTCATGCCTTTTCCTGCCTCTCTGATCTTTTC
TAGGTCTCTGGCCTATCAGGAGGACAACTGGTGCTGCAATAGAAGCCAGTGGCTAAGTCT
[C,T]
GTGTATGGCGTGGTTAAGGTTGCAGCCTCTCACCTCTGCCTTCCTCCATTTTGGGCTGGT
TACCTTTGTGCTCTTCCTGAATGGTCTTCGAGCAGAGGCTGGTGGCTCAGGGGACGTGCC
AAGCACAGGGCAGAACAATGAGTCCTGTTCAGGGTCATCGGACTGCAAGGAGGGTGTCAT
CCTGCCAATCTGGTACCCGGAGAACCCTTCCCTTGGGGACAAGATTGCCAGGGTCATTGT
CTATTTTGTGGCCCTGATATACATGTTCCTTGGGGTGTCCATCATTGCTGACCGCTTCAT

2381 CCTGAATGGTCTTCGAGCAGAGGCTGGTGGCTCAGGGGACGTGCCAAGCACAGGGCAGAA
CAATGAGTCCTGTTCAGGGTCATCGGACTGCAAGGAGGGTGTCATCCTGCCAATCTGGTA
CCCGGAGAACCCTTCCCTTGGGGACAAGATTGCCAGGGTCATTGTCTATTTTGTGGCCCT
GATATACATGTTCCTTGGGGTGTCCATCATTGCTGACCGCTTCATGGCATCTATTGAAGT
CATCACCTCTCAAGAGAGGGAGGTGACAATTAAGAAACCCAATGGAGAAACCAGCACAAC
[A,C]
ACTATTCGGGTCTGGAATGAAACTGTCTCCAACCTGACCCTTATGGCCCTGGGTTCCTCT
GCTCCTGAGATACTCCTCTCTTTAATTGAGGTGTGTGGTCATGGGTTCATTGCTGGTGAT
CTGGGACCTTCTACCATTGTAGGGAGTGCAGCCTTCAACATGTTCATCATCATTGGCATC
TGTGTCTACGTGATCCCAGACGGAGAGACTCGCAAGATCAAGCATCTACGAGTCTTCTTC
ATCACCGCTGCTTGGAGTATCTTTGCCTACATCTGGCTCTATATGATTCTGGCAGTCTTC

5165 TTCCTCTGAATGACTGAACATATCCACAAATAATAAGCGTGGCAGGAGATGGTGTGAAGA
GTAAAAGGAGCATATAGGAAGTTGTGTGTGTGGGGTGTCTGTTTCAAGAACCTGCTAATT
ATACCTTCAGTAAGAAATGAAGCCATACAACCTCTAGAAGAGGAGGAGGAAGGAACTCAT
GGAAAAGTGGGGAGCCATAGAAGCTAGGGAGAGGTGTCCTAGGAGTGCTTCTGCCCAGGT
CCAGCCATGAGACAGAGCTCAAAAAGAGCTGGGCACTGCTGGTGACAGAACTGAGTGACC
[C,T]
GGGGGATCCTGCATCTGTTCTTACTCAATCCCTTCTTAATAATGTGACTTGGGGCAGGTC
ATTTATTGGTTCTGGAACTTAACTTTCTGATATGCAAACTGGGAATAACAATACTTTCCT
TGCCTGGAGGCAAGGTCAGTCCTTTTTGCAGTTCCTTCCAGCTCTAAGATTTTCTGAACC
ATAGACATAAGCACTCAGTGTAGGTCATATTCGCACTTGCCAAAAATGGATCAGGGAATA
TTGTCTCCTGAAGGGAAATGGCCATTGACAAATTGATTTATTAGAGCTCTGTTTAGTCAT

5402 GGTCCAGCCATGAGACAGAGCTCAAAAAGAGCTGGGCACTGCTGGTGACAGAACTGAGTG
ACCCGGGGGATCCTGCATCTGTTCTTACTCAATCCCTTCTTAATAATGTGACTTGGGGCA
GGTCATTTATTGGTTCTGGAACTTAACTTTCTGATATGCAAACTGGGAATAACAATACTT
TCCTTGCCTGGAGGCAAGGTCAGTCCTTTTTGCAGTTCCTTCCAGCTCTAAGATTTTCTG
AACCATAGACATAAGCACTCAGTGTAGGTCATATTCGCACTTGCCAAAAATGGATCAGGG
[A,G]
ATATTGTCTCCTGAAGGGAAATGGCCATTGACAAATTGATTTATTAGAGCTCTGTTTAGT
CATTTTGCTGGGAAGGATAATCATTTGTTAACGTAAGTAGAAACCTGTGCCTTCTGGAGA
ATACTATCCATTTATATGTACTCTGGGGAGAGTGTTTATACATACAAATGAAGGACAGGG
CTTCACTGGGAAAACAAACTCCATGGAATTTCACATGATTATCGCGATGTCAGTGTGGAA
GAAGATATGGTAAGGCATTAAATGACATTAAGACCACAAAATTTGCCATAATTTGACGGA

6794 CTCATAAAATATTAGAGCTAGAAAGGACCTTAGAATATCTTCTGCAGTCATGGTTCTTAA
ATTTTAATGTGTTGCTCAATCATCCAGGGATCTCACTGAAGGGCAGATTAGGATCCAGGA
GGTCTAGGGGAGGGATTGAGATTCCGCATTTCTAACAAGTTCTGGATGCTGCGGGCCCCA
ACTTAGAGGTGAAAGGTTCTGAAGCTCTTGACCAAACCAGGAGACCCAGCAAAGAAGTGG

TTTTTCAGACAACTTGCTTAATTGAATAATGATTGTTTGCTCTTTAATTCCAACTTTCAA
[T,C]
GCCAATTTAGCAAGAACCAGAGGCTGTGCTAATTGCCACACCAGTCTGGAAACCGAAATG
GATAGCTTCAGGGTACTTGGACAAAGTTGGAACATCTGCTTTCTAATCTCTCCCTCTTTG
TATAGCTTTATTTGCCTACCAAGCCTGGTAGTATTGAAAATCTGCCCTCACTATACTCCC
CTAAATATAATCAAGTTGAGGCCAGGCCTGTGCTCTATCAATAATATAGGATCCACGAAT
TCACATGTTTGGTTTTATGCTTTACTTCTTCAAAGGTGCTTTTAGCAGCATGGAAGAATG

9883 GTCAAAGAATATGTCAAAGCATGACATATTCCAACTCCAGGATCCATAAAACACCCCAAG
TTCTGTGGAGACCCTATCACATCTGCAAAACTCTCCAGGAAGTCCAGAGCCCTCCTGGTT
AATTTGTTTTAGGGACTAGGCATGCGGTATCCCCTGACAACACTGGATCAGCAATTCTCC
TACCTAAGTCAGTCCCACACCATGTGCAGCAGAGTATCCAGTGCCCCTGCCCTGGTCTGC
TCACATTGGTTTGCTCTCCAGAATAATAATTCCTCAATATCCACAAGAGATTGATTCCAG
[A,G]
ACTACTCCGAGGATACCAAAAATCCTCAGATGCTCAAGTACCTGGTATAAAATGGCACAG
TATTTGGCATATGACCTAGGCATATTCTCTCCCATATACTTTATTTATTTATTTATTTCG
GGACAGAATCTCATTCTGTCGCCCAGGCTGTCACTCGCTTATTGCAACCTCTGCCTCCCA
GGTTCAAGCAATTCTCCTGCCTCAGCCTCCTAAGTAGCTGGGACTACAGACGCATGTCAC
CACGCCTGGCTACTTTTTGTATTTTTAGTAGAGACAGAGTTTCACCATGTTGGCCAGGCT

10210 CAGATGCTCAAGTACCTGGTATAAAATGGCACAGTATTTGGCATATGACCTAGGCATATT
CTCTCCCATATACTTTATTTATTTATTTATTTCGGGACAGAATCTCATTCTGTCGCCCAG
GCTGTCACTCGCTTATTGCAACCTCTGCCTCCCAGGTTCAAGCAATTCTCCTGCCTCAGC
CTCCTAAGTAGCTGGGACTACAGACGCATGTCACCACGCCTGGCTACTTTTTGTATTTTT
AGTAGAGACAGAGTTTCACCATGTTGGCCAGGCTGGTCTCAAACACCTGACCTCAAGTGA
[T,C]
CCGCCCACCTTGGCCTCCCAAAAAGCTGGGATTACAGGCGTGAGCTACCACGTCCAGCCC
CCCATATACTTTAAATCATCTCTAGATTACTTATAATACCTAATACAATGTAAATGTTAT
ATAGTTGTTTTAATGTATTGCTTTTTTATTTGTATTGTTTTTTATTGCTGTATTATCCT
TTTTTATGTTTTATTTTTTCAAATATTTTCTACCCGTGGCACCCACAGTTGGTTGGTGGA
ACCTGCGGTTGGTGGAGCCCATGGATGTGAAGGGCTGATAGTATGAGAAAACTCAGAGGT

12220 ACATCCAAATAGTAACTTAATATTCCAAATATGGCTGCAAAACAAATTGTCGATTATGGA
TGACTACTACTGCCATCTCTCCATACCAGTCCATCTTCTGCCAGGCTGTTTGGTCTTGAT
TTGTCGACCTTTTAGGTTTCTCCCCATGTATTCCACATGACCTTCACCAACCCCACTTCT
ATCTCCAAACGTCTTTCTGAGTTGTGGGGATGCAGATGTATTCTGCCACCATCACAAGGG
CTAACCGAGCCCTGGCTGCGGATCTTCATTGTTGTTCACATTATTTCCATTCTTACACCC
[T,G]
ACTTCATGTTTGTACACTATTTTCTTACATTTGCTGTCTCTTCTAAACATTCTTTGCTGC
ATCCACTTTTTCTCTATTTGTGCTCTAGGTGCTGCAGAGGCTAATGCTGGGTTTCCTTTC
ATTCCTCCTTGCACTCAGCACCTCCCTTCTCAATTCCTTTTGCCATGTCTCCACTTTAAA
TCTTAACCTACTCCAGATAGTCTTTTCCTTCACACTATTGGCATCTGTGCTTGGGTTGCT
TTCAGTCTATTCTCTGATCTATGATTTCTTTGCATGATCAAGAAGGTGCCATGAAAGGAT

13842 TCACTTTCAAAGCCTCTTTCTGGGTTTGGATTTCCAGAGCAGCCTGTGCTGTAAAGCAAG
ACAGAAAGCTTCCCTGCCATTCATGCCTGCCAGGGATAGAATGACAGTACTCCTGAGGCT
CTCCCTCCCCACCCCTCCCCTGCTGGACAGCTGATCTGCTGGACTCAGCCAGAGCCAGCA
GGCACCCCCCTCTTTATCCTAGGAGCTGCAAACTTGATGCCTTTCCAGGAAATCCCCAGAA
GCTGGAGTATCCTCATCTACATGTGGCACAGTGTATGGTTGTGTCAGGTGCTCATGTCCC
[G,A]
TTGCATAGGACTGGGGTGGAAAATAGGGACCGTCCTTTTGTGTCAGCTCCAGTCAATGAG
TAGTGGCCATCCAGGGGGCCATCTTGGAAAGGACTTGTGAGGCTGTATCTGCGCTCAGTT
GTAGATGTGAGAAGAAAAGGCCAAATATCTGCCAATCCTAGTCCTGGGATTCAAGATAGA
AAGAACTGCATGGAGTGAAGAAACTAGGAGTCTCCATTTCACTGAGATGCATAAGAATGA
AATTATTGTCACTATTTCTTCAATACTGGGCCAATCCTAATAAGAAAACCCTTTTTGAGT

14200 GAGTAGTGGCCATCCAGGGGGCCATCTTGGAAAGGACTTGTGAGGCTGTATCTGCGCTCA
GTTGTAGATGTGAGAAGAAAAGGCCAAATATCTGCCAATCCTAGTCCTGGGATTCAAGAT
AGAAAGAACTGCATGGAGTGAAGAAACTAGGAGTCTCCATTTCACTGAGATGCATAAGAA
TGAAATTATTGTCACTATTTCTTCAATACTGGGCCAATCCTAATAAGAAAACCCTTTTTG
AGTCTCTCTTTTCTTTATCCTACATATAACACAGAAGCTTTTTCTATTCCCTGGATGAAC
[C,A]
CACAGGGACAGAAATTCTTGTTGGACAGGTGAAGCAGATAATTTCTTTATCAGACTAGAA
TCTTCCAGAAGCACTGCTAACCTAGTGAGTTTTGTACTCTAGACAGGTGGTTCTCAAGCC
AGCTCCCCACCGCAGGCCTTTTTCATGGTCTGCCCCTCCCTGTGGAACCCATGTTTTAGG
TTATTAGCTGATAATTGGATTTCTATTTTTTCTCATAAAATACAGCAAAAGATAGCTAGT
GATATTATGATGAGTTAATGTAATTATAGCCAAAGCAGAGAGAAACAACATTTTAATTAA

15878 TGTGTCAAAATATCACTCTGTATCCATACATATGTATAATTATTATGTGTCAACTAAAAA
TAAAAGGAAAAAAATCATTTCAGTGTATTTACAAAACATATGTAACCATTAAGAATAATG
TTTTAAATTATATCTAAGGGTGTGATAAAATTACAGTATAAGATTGTGCTTGAAAAAGTG
CAATAAGAAGTAAATATGTACAGATGAGAAAAAGTGCAAAGAACTAAGTCCTAAGCAGAC
TATACCTTTCCTACTGCATGGTACTTCTCTGGCCTTTTGCTTTGAAAGATTTTGCACCCA
[G,T]
CATGGCAAGTGGTTAGCAGAGGCAGCCATTCTCACTTGTGCGTTGGCTTTGGGAGCCATA
TATGTTGTTCAGCTGGGTGTGGAGTGGAAAGGCTGCATGTTGTATTAATGCATTGTTAAG
AACCTCTAAGAGTGATTTCTTTTGGGAAGTGAGACTGACGGTCCGAATGGTGGAAAGACA

ACTTTTAATCTTTTACTTTACACTTTGTGCACTTTTAAATGTTTAACATGAGCATGCATT
TCTTTAATAATAAAAATACAAAAAAATTTTAGCCCTAGATCTTCTGATTTTAAACTGCAT

16030 ACAGTATAAGATTGTGCTTGAAAAAGTGCAATAAGAAGTAAATATGTACAGATGAGAAAA
AGTGCAAAGAACTAAGTCCTAAGCAGACTATACCTTTCCTACTGCATGGTACTTCTCTGG
CCTTTTGCTTTGAAAGATTTTGCACCCAGCATGGCAAGTGGTTAGCAGAGGCAGCCATTC
TCACTTGTGCGTTGGCTTTGGGAGCCATATATGTTGTTCAGCTGGGTGTGGAGTGGAAAG
GCTGCATGTTGTATTAATGCATTGTTAAGAACCTCTAAGAGTGATTTCTTTTGGGAAGTG
[A,G]
GACTGACGGTCCGAATGGTGGAAAGACAACTTTTAATCTTTTACTTTACACTTTGTGCAC
TTTTAAATGTTTAACATGAGCATGCATTTCTTTAATAATAAAAATACAAAAAAATTTTAG
CCCTAGATCTTCTGATTTTAAACTGCATATTCTTTCTATTGTGTTACATATTTTAGCATG
AGAATAAGGTTATGAAGCTGGAAGTAGCAGGCTCCCTTTTCCTCATATGTAGGAAGTTAA
GAATGCATTCTACGTTTCTTCTTTAAGGAGTTGGCTTCTTTCCTTTTAACATAGGGGTAA

16292 TGTTAAGAACCTCTAAGAGTGATTTCTTTTGGGAAGTGAGACTGACGGTCCGAATGGTGG
AAAGACAACTTTTAATCTTTTACTTTACACTTTGTGCACTTTTAAATGTTTAACATGAGC
ATGCATTTCTTTAATAATAAAAATACAAAAAAATTTTAGCCCTAGATCTTCTGATTTTAA
ACTGCATATTCTTTCTATTGTGTTACATATTTTAGCATGAGAATAAGGTTATGAAGCTGG
AAGTAGCAGGCTCCCTTTTCCTCATATGTAGGAAGTTAAGAATGCATTCTACGTTTCTTC
[T,C]
TTAAGGAGTTGGCTTCTTTCCTTTTAACATAGGGGTAACTGGGCCCAGGGAGTTTGGCAA
GGGCCAAATAAAGTCCTTAATGCCCAGCTCAGAAATCTGGATTCACCATCCTTGACTGCT
GGCTCCAACCCACCCTCACCTGAGCTGGTCTGCAGAGGATTCTTGTTTGTGTCACTTCAT
CACCAGCAACTACCGACAGATGATGCTTTGGCCTGCTGCCTGGGTAACAGGGCGAGGCTG
GCTCAGGACCATGTTTTCAGATCAGGGGACCTCCTTTGATGCCATGTCCATGGTGTCCGA

16506 GCATGAGAATAAGGTTATGAAGCTGGAAGTAGCAGGCTCCCTTTTCCTCATATGTAGGAA
GTTAAGAATGCATTCTACGTTTCTTCTTTAAGGAGTTGGCTTCTTTCCTTTTAACATAGG
GGTAACTGGGCCCAGGGAGTTTGGCAAGGGCCAAATAAAGTCCTTAATGCCCAGCTCAGA
AATCTGGATTCACCATCCTTGACTGCTGGCTCCAACCCACCCTCACCTGAGCTGGTCTGC
AGAGGATTCTTGTTTGTGTCACTTCATCACCAGCAACTACCGACAGATGATGCTTTGGCC
[T,G]
GCTGCCTGGGTAACAGGGCGAGGCTGGCTCAGGACCATGTTTTCAGATCAGGGGACCTCC
TTTGATGCCATGTCCATGGTGTCCGAGGGCAGCCAGGATCAAGGGCTAGACGGGGCAGTG
ATGAGATGAGAGCAGGAGGGGCTCAGCTGCAGCCCCAGGAGAGCCTATGCCAGCCCTGTT
GACCAAGGAGGACAGAAGCAACAGGAGAGCGGAGGCAGAGGGGTGAGTGTCTATCGCTCA
ATGTATAATCGGCAGACATTTGGGGAGCTCATACTGTGGGCTAAGCACAGGGAAGAAAGG

17953 GATTGGACGCAGTTCTGCACAGCACTTTTCCGAATGCCTCTGAAATGAGTCCTCACTGAC
AGAACGGGCCCACTCTGGGGGAACTGAGGGCTCTCTTGGTCCTGCACTGCTCTTTGCCAT
ACAGATCTGTCTGCCCAGGATTTTTCTTGGGTGTGTAGGAGGCTGAGAGAGCTCCCCTTT
CTTCTCATGGCTAAATCCCTTGGTCTTTCCAGCCCTCCTGGGGGTTAGAAGGGAGAGGGA
AAAAAAAAAAGACTGAACTTGTTGTTGTTGTTTTTGTTGTTGTTGTTGTTTGCCTGTTTT
[C,A]
TATGTTGTCTTGTGGGGAGAGGGTATAAGATTGATTGACAGAGTGGCACACTTCCCCTGC
AAATTCATCATTTGAATTTCTCAGGTAAGATGTTCACATTTCTCTGTTAAGATGCTCCAA
TTTCTCTGGTTAAGATTTCTCTGGTAAGATGCTCATGAATTGGTGGAGGTGTTGGCGGGA
TGTGGGAAGTGTGCCTGCTCTTTCTGAGTTTTGGGGGAAGTTGCCTTAATTCTCTGCATG
ACTTTCTTTGCTCCTTTGGGCTTCATTTCTGTGCAATGTAGTCTGACATGAATACTGCTC

23832 TAGGCAACAGCATTATAACTCCTGCCTTCACAAAGCTTATCTAACACACACATTTCCTCC
TCAGGCACATCCCAGCCTTCTTGCACTTAGGATTCAGCAGTATGCTTAAGGGCCATTTTC
AACAGCAAACTCATCAGCGCAAACACAAACATGTGAAAAACGTAGCACTAAAGAGACTGC
AAAAAGGACACTGGCTTACAGCATGGAAGCTGGAAGGAGAAGGCAGAGAATCACCTTGTT
CCACTTCAGCTATGAATATGCAGTCAGGCCACCCAGTCATTCAAATTTTATAAATATACT
[C,G]
TAATATATATATAAATACCAGGCAGGGTTATTTTTTTCCTCAAGTCATTTTTCTAATTTT
TTTTAAATGAATAGATAGAAGAGCTGAAGTAAGGGTCAGGAGCAAGAGCTCTGCTTCCTT
TTCCCTTGCTGGGCTTCGTTAGAGAGCCATCATCTCCTCAATATGTCTCCCAACTCTTCT
AGGCATTGGATGAGTTTGCTGCAGATACGAAACCCAACTTTGCCAGTCACTTCATACTAA
CAGGTGAAATGTAGTGGAGGAGCCTTTTGAAGACAGGGACTCAGCCCCCCATTAGCCTCA

25001 GGGCAGATGAGTTTATACGTTTCTTTCATGTCCCCTTCCTCCCACATAGACTTTTATTTC
CCCAAAGGAAAACAGAAAACAATGATCTGTTTGACAGTGTTGCTATCATTGGGCATCAAA
CCTATCATCTAAGGGGAATCCCCCTGTATAATCAGTCAGCCAAATGGAGCAGGACCCTGT
GTTTTGTAGCTGATACAACAGGGCAGCATCTCTAGTGAGGGGGCCAGGGCTTCTATTTCC
TTCATTAAAAAATGAAACAGCAGACCTGATTCCATATTTAGAGATTACACTTAGTTGCCA
[C,A]
TGTGGGTGTGCAGGCACCAACCAAACCCAGTTGGCACCGTTGTCTTTTCTCTGCAATGAT
GTATTGAATTTAATAATGGAGGTATATGAAATTCAGAGTGATTGGAACTGAAGGTTTAGG
GGCTTTGTGTAAAATTGATATGTAAGGGATTTGGAAGTAGGTGAGGGATTCTTCCCCAAT
ACTTATTCAATTTTGGAGTCAAATAACCAAGCATTTACAAATAGCCAAAAAAGAAATTGA
AAGAGGGTTTAATCCAATAAATTTTCATGCCTCATATGAACCACATCTTATAATAAGAAT

25141 CCCCTGTATAATCAGTCAGCCAAATGGAGCAGGACCCTGTGTTTTGTAGCTGATACAACA
GGGCAGCATCTCTAGTGAGGGGGCCAGGGCTTCTATTTCCTTCATTAAAAAATGAAACAG

CAGACCTGATTCCATATTTAGAGATTACACTTAGTTGCCACTGTGGGTGTGCAGGCACCA
ACCAAACCCAGTTGGCACCGTTGTCTTTTCTCTGCAATGATGTATTGAATTTAATAATGG
AGGTATATGAAATTCAGAGTGATTGGAACTGAAGGTTTAGGGGCTTTGTGTAAAATTGAT
[A,G]
TGTAAGGGATTTGGAAGTAGGTGAGGGATTCTTCCCCAATACTTATTCAATTTTGGAGTC
AAATAACCAAGCATTTACAAATAGCCAAAAAAGAAATTGAAAGAGGGTTTAATCCAATAA
ATTTTCATGCCTCATATGAACCACATCTTATAATAAGAATTATGCTTTTTCATTTCATAC
TCAGTTAACAAATATGATTTGTGAGCACCTGGTAAGTTCAGGGCACTAGGCTGAAAGGGG
TTACCAAATGTCTTCATTTAACAAAGTCCAGCTGAGCTCTTACAGGTACCAGAACTGTGC

25191 TGATACAACAGGGCAGCATCTCTAGTGAGGGGGCCAGGGCTTCTATTTCCTTCATTAAAA
AATGAAACAGCAGACCTGATTCCATATTTAGAGATTACACTTAGTTGCCACTGTGGGTGT
GCAGGCACCAACCAAACCCAGTTGGCACCGTTGTCTTTTCTCTGCAATGATGTATTGAAT
TTAATAATGGAGGTATATGAAATTCAGAGTGATTGGAACTGAAGGTTTAGGGGCTTTGTG
TAAAATTGATATGTAAGGGATTTGGAAGTAGGTGAGGGATTCTTCCCCAATACTTATTCA
[A,G]
TTTTGGAGTCAAATAACCAAGCATTTACAAATAGCCAAAAAAGAAATTGAAAGAGGGTTT
AATCCAATAAATTTTCATGCCTCATATGAACCACATCTTATAATAAGAATTATGCTTTTT
CATTTCATACTCAGTTAACAAATATGATTTGTGAGCACCTGGTAAGTTCAGGGCACTAGG
CTGAAAGGGGTTACCAAATGTCTTCATTTAACAAAGTCCAGCTGAGCTCTTACAGGTACC
AGAACTGTGCCTGGGCTGTCATATGAAGATGAATGTAAGAGTGTGTCAGGCCTTCAAGAG

26147 GCATGATCTCTGCTCATTGCAACCTCTGCCTCCAGGTTCAAGCGATTCTCCTGCCTCGGC
CTCCTGAGTAGCTGGGATTACAGGCGTGTGCCACCATACCCAGCTGATTTTTGTATTTCT
AGTAGAGATGGGGTTTTGCCCTGTTGGCCAAGCTGGTCTCAAACTCCTGACCTCAAGTGA
TCTACTCGCCTTGGCCTTCCAAAGTGCTGGGATTACAGGCATGAGCACTGTGCCTGGCCT
TTTTTTTTTTTTTTTTAAAAAAAAAAAAAAAAAAAAACAGGAAGTTTTCGTTAGTTTTTTT
[-,A,G]
TTTGTTTTACTTCCCATAAAAACTCTTTGTGTCACATGGAGGTGAATGGAAAGAGAGGCT
GTGGCAACAGACGGGAGACTTTTCTGATATCAGAACCCAGTCCCATAGACCAGAATGTAT
GCTTTCAATCCACGTTGTCTGGGTCCATCCTATTGAGTGCCCTGCCCCCACAGCGGGGTA
TGGAGAAGAGTCAGACACAGCCCCAGTCCTCACGTAGCTCACAATCCAGTGGAGGAGACG
GACTCAGAAACAGATAGAGATGAAGCCATGAGATCAGTACTGTCCGAGGCCATGGCCACG

27400 TAAACTTTACAAATCCTTAATTTGTAAAATGTGGGCAATGATAGTACCTCCTCACAGGAT
TATTACGAGGTTTACACGGAATACTCTCAGCTCATAATAAGCACTTGCACAGGCCTCATG
GGCTAGGCCCTCAAAACTTAACGCATCTACAGGCAACAGCCATATGAAAGGAATTTTATA
CCACCAAGTCAAAAAATCTGTGAGCACTGCTCAGAAGCAAAAGCCTGTCTCCAACAGCGC
TCATTTAAGGGGTGGGCGAGCTACAGAGAGAAGAATGAGCCCCCACAGGGTAAGCTGGGG
[A,G]
AAGCTGGGGACAGAATGAGACTCAGGAAATCACTTGAATATTGATTATATTTGTGCTCAA
TAATAAAATAACGAAATGAGTACAGCCCTAGACCTAAACATTGTGGGTGAGGCAAAGGCA
ATGCGTTAATTTTGCATCCACTGAGGAAAAACTCTAAAACGGTGACTTCTTTTTTAAGGG
ACCAGAAGAATCTAGATTATATTTAGTCTAAGTCAATACATACGACAGAACCTTGCCCTC
TAGACTTGATAAGAAAGAAGTAAAATAAGAGAAAGAATAAAAAAACCCTTCCACCAAAATA

27401 AAACTTTACAAATCCTTAATTTGTAAAATGTGGGCAATGATAGTACCTCCTCACAGGATT
ATTACGAGGTTTACACGGAATACTCTCAGCTCATAATAAGCACTTGCACAGGCCTCATGG
GCTAGGCCCTCAAAACTTAACGCATCTACAGGCAACAGCCATATGAAAGGAATTTTATAC
CACCAAGTCAAAAAATCTGTGAGCACTGCTCAGAAGCAAAAGCCTGTCTCCAACAGCGCT
CATTTAAGGGGTGGGCGAGCTACAGAGAGAAGAATGAGCCCCCACAGGGTAAGCTGGGGA
[A,T]
AGCTGGGGACAGAATGAGACTCAGGAAATCACTTGAATATTGATTATATTTGTGCTCAAT
AATAAAATAACGAAATGAGTACAGCCCTAGACCTAAACATTGTGGGTGAGGCAAAGGCAA
TGCGTTAATTTTGCATCCACTGAGGAAAAACTCTAAAACGGTGACTTCTTTTTTAAGGGA
CCAGAAGAATCTAGATTATATTTAGTCTAAGTCAATACATACGACAGAACCTTGCCCTCT
AGACTTGATAAGAAAGAAGTAAAATAAGAGAAAGAATAAAAAAACCCTTCCACCAAAATAC

29278 ATACACTTCAGCAAGTCACCTAACCTGCAAATTTCAAGCATGTGAATCTTGGATCTTTCA
TGTGCTAGCTGTGAGACTTTGAGAAATGTATTTAATGTCTCTTTGCTTCCTTTTCTACCC
ACACAATGGGTATAATAATGTCTACCATATATCTTTGCAGCAAGGTCTAAATGGGGTGAT
ACATGCTGAATACATTTCCAACAGAGTCTGTGCAATGATAAGCTCTTTCCAAATGTTAGT
TAAAGCTAACCAACTAACCCACCAACAAACCAACCTCTTAGCCAGGACTGATGGAAGGAG
[C,T]
CTGTGAGAGAATGCATTTAAAACACTTGGCACCATGCCTGACAAGAGTAAGTACTCGATA
AATCAGTTATTGTTATTATCGCATCGGTATTATGACCATTATCCTCTTCTCTATAGGCTT
CAGGTTTTCCTGTCTTTTTATCACAGCAGTATTCCAGCAGAAGCCTTTGATTTAACTAAG
TCTCTACTGTGTGTGTGGCTAGATGCTATAAAGCATCCAGAGAAGTGAGAATTTGGTCCT
GCTTTTAAGTAGCTTATAGTCTAATTAGGGGGAAGTAATCAGATAGAAAGGAAACTAACA

31437 ACTTGGCTTTGCCGGGGTAAGAGGGGGCACTTCTCTCCTTTCCCTCATGAAAGGAGGGAG
AGAAGCCAAAAATCTCCCTACTAGTCAACAACTCAGGCACCCCTCCTTCTCTCCTCTATT
TTATAGACTGGGAAGGGAGTGATGGTTGTTGGAGGTGGCAGAGCCAGTTCAGCTGCCTTT
TGTGAAGTCCTGAAGGAGGTGTCTATCCTCAACTGCTGGCTTCTGTCCTTAAGCCTGGGG
AGAATTAAGTCCTCTTTGCCTCAGTTTGGCACTCCAATTGCCAACATTGGGACAGCAGGA
[A,G]
AAGTTCCATCCAACATCCCATTAAATATGTAATGTGTATTAGCACAGCGCCTGGCACTGG

GCAGGTATTTTCTAAGTGATAGCCAATGCGAAGCCTACTTTATTATTTTCCTCTTTGCTT
AACCTACAAGGTGTCTAAGACCATTTGTTTGTCCACACATAGTAAGATAAACAGCACTGA
GACTGTGGTCCTTTCTGCCCTGTGTCCTTATCCCACCTGGGAATCTGGAAAGCCAAGCCT
AGACACACTCGTTCCACAAATGTTTACTGAAGCTTGTTCTATTCAAAGCACTGTACAGCT

31857 TAACCTACAAGGTGTCTAAGACCATTTGTTTGTCCACACATAGTAAGATAAACAGCACTG
AGACTGTGGTCCTTTCTGCCCTGTGTCCTTATCCCACCTGGGAATCTGGAAAGCCAAGCC
TAGACACACTCGTTCCACAAATGTTTACTGAAGCTTGTTCTATTCAAAGCACTGTACAGC
TACAAAGACCATCTTTTCTGAACTCCAAACCAGGCCACATGGTTGGAATAACTTCAAGTA
TGGAGACCAAGAGAAAAGGTGGTTGTTGTCAGCAAAGCTCTGAGTCCACACCTTCCAGGA
[A,G]
CTTATAGTTGATGCAATGGTGGGAGAAGTCTGAACCTGGATTCAATCTGCTTGATTCCGA
TGAATGGTGCAGTAGGCAGAGCCATGAGTTCAGAGCAGGAAGAAACCACTGGTTCAAAGA
AGCATCTGTCACATCGAAGCTGCTTTATAGTCTGTTGGGAAGCATGCATAATAATTTATT
CTTTCTTTCTTTCCTTTGGTCAACAAAGATTTCTTGAGTCCCTACTATGTGCCAGGTACT
CTTCTAGGTACTGAAGATGCAGCAGTGAACAAAGAAGATACAATCCCTGCCCAGCGGAGC

33155 ACAGTGCTGAATTTTCACAAATTGCGAATTAGGAAATTGTTGCTCATTTTACAATTTGGT
TTCCCTCAGGATTCCTTTTAAGTAGCCAGCTACCCCAGTACTTTTGAAATATGACTTGCT
TATAAAAATTTGATAGGCTTGGCACGGTGGCTCACACCTGTAATCCCAGCACTTTGGGAG
GCCGATGTGGGGTGGATCACGAGGTCAGGAGTTCAAGACCAACATGGTGAAACCCTGTCC
CTACTAAAAATACAAAAACTAGCCAGGCATGGTGGCACATGCCTGTAATTCCAGCTGCTC
[G,A]
GGAGGCCAGGCAGCTAGGCAGGAGAATCACTTGAACCCAGGAGATGGAGGTTGCAGTGAG
CCAAGATCATGCCACTGCACTCCATCCTGGGTGACAGAGCAAGACTTCATCTCAAAAAAA
AAAAAAAGATATATAAACAAGTTTTTATAATATTCTCAATATGAACTAGTAGAAAAAAAG
CATGTGTTTTTAGGTCTTAGAGGCCTGGTTCCCAGTTTTATCTCTGACTCTAATGAGGTA
TAGTATTACCTACATTGATTAGCCCTTCTATACTTCATAGGAGATGCTCCAAGACTGCTA

39487 CACTTTGCTCCATCCCTTGGCCTTCTGCAGTCCAAGCTCCATTCTGAGATCATCCAAGGC
TTCTCTTCTGTGTTGATCCTTGGCCTTCTTGGAGTCTCTTTCTCCCATGTTCTCCACAAC
AGAGCATTCTCCTGACTGTTTTCATTCTGCATCTCACTCTTTCATCAGTATCTTTTTCTC
TACCATGCCCCATAAATTTGGGTGCTCCTGAGGGTCCTGTCCTTGTCCCCTGCTTTCTTG
TTGTACAACCTCCTTGATCTACTTCATCTACTCAAGTTTGGTCCACAATTTCTATATTGT
[G,C]
AAGATTCAAATCTGCATCTCTAGCCATATATCCATTTGCCTGCTAGGCATTTCTACCTGA
ATATTTTATAGGCATGCCAGTGGCTCTTACTCTATGGCTCTTACTCTAAGTCTAGACTAC
AGCAGAAAGCAATGCTCTTTTTATTAAGGCATAGTGCCTCTTTCAGAATAATTTACAGCA
TACAACCAGGCCTGCTGTGCAGCATTACAATTTGTCATTAAAACTCCATTCCTCTTGCCA
GAGTAAATGAGCCATTTACAGCCAGGGCGCCAAGATGGACTGTTGTTATTTTTTCTGCCT

41449 TCAGATTCCAGGACACCAAGTTTTCTGTGGGAGCTTCCCTAGGAATATAACTAAGGAATT
TAAATCAGGTTCAGCTCATGCTGTTACACTCTCTTCCTCCACTCAGGCATTGGGTGTGGC
TTTTCCAAGCTTGAGAAGGGTGTGATCTGAGATGGGCTTGGGTATAGAGGGGAATTATAT
TTAGGTCTACCCTGTATAGGAAAAAGTGCCTTCCCAAAGTCTCCCTGGCCTAAAGTATAA
GAGATATGTGTTGGGATTTAGACCCAGAGCCCAAGCCAATAATGGGACCCCCTTCTCACA
[T,C]
GTGGCTACCTCCTGCTATCACCACAACAGCTATCATACCCATAACTACAACAGAGGCCAA
TTAACGTGGTGATAATTGACAAATGTCAAGACATCCTACATTGAGGCACACTGTGCGTTT
TGCGTGAGCTTTTAAATTGGTAGGGAAGGAAAACTTTTATACCTACACCTATCATGGAAG
GCAGAAGGTAAGAGCTAAAATAAAGGTATGCCAAGAACAAAGGCAGGAAAGAAGGGTTTT
AACAACTTGAGGCCTGATCCATTGATTAGTGAAGAGGAAACATGTTCAAAAACCACTCTA

42420 GGGTGACATGATAGATCTGTATTCTAGAAAAGTTAGTTTTGCAGCAGTTGTGTCCATTGA
AAGGGACAGGATAAGGGAGATAGATAAGAAGACATGCTATGATGATAACTAGATTTGGAT
ACCAAGTGGTATGGTGGAAAGGAATGAGAGAACAGGGTCACAGATGAATGACTGCCCAAT
TTCAATCCATCATAACAGGATGTATAGGATTGCCCTTAAGTAAGATGGGGAATCCAAAAA
CGAGGAACAAGTTTGTAAGGTTTTGGGGGCCAATGATGAATTCCATTTGGGACATGTTGC
[T,C]
TTGGATATACCAATGGGACATTCATGTGAAAATGATCTCGGCAATCCTATCCTGGAATTC
AGGATAGGATCAGAATGAGGGACACAGTTTATAAGGTAAACAGAATGGAGGTGATATAGA
AGATAAGGGCATAGATGAGCTTACCAAAGGGGAGAGTTTAGAATGAAAAGAAAAGACCAA
AGGCTAAGCCTGTGCTATTCTTTCTCCTCACAATACGCTTCAGACCTGGGCACAAACCAT
CAGTGAGTGTCATGATAACACTACTGTGGGCAAATCCCCCCTCTATAAGGGCCTGATTTC

43256 AGAAAAACAATTAGAATGGAGAGCTAACTCTTTGGAAATGGTCAAAGAACACGGGTCTAC
AAAACCGTCAATAAAGCGCTAAGATGCCTGGGCGGGGTCAAAAAGTCTACCTGGGCGGGG
TCAAAAAGTCTACCTGCTCAGCATATGGGGCCCAGACATCTGACCTTTACCAACTCCACA
ATAACCACTTCATCTATGGATCCAGTCTTGGTATCACCTAGTCGCTGTTTTCAAGTAACA
GAATATTTGGTTCTCAATGGTAGGTGACTGGAATACAGCTTACTTTCTCCCACCCCTACC
[G,C]
CCAATCCTTTCTGCCCCCTTATAGTTTAATTTGCTTGTAAATTACTTGGGAATACATTTG
GGAGCCATTATAGGGAAATAGAAGGCAGACATGATGAACAGAATGCAGGGTGTTTTTTAT
TACTTCACATTGTGCTCAACAATTAGGAGGAATTCTAGAAGCCCCTCCCAGTGGCCAGGA
ATTGGTCATAGCATGAATAAACTCAATATAGGTTGAGTATTCCTTACCCAAAATGCTTGA
TACCAGAAGTGTTTTTGGATTTTGGATTTTTTTTTGAATATTTGCATTATATACTTACC

43967 GGGTTTTGGGATTAGGGATACTCAACCAGTGGTAGGTTTGGGATGATATCAGCATGCTAA
GGTCAAAGAGACCTAGCTGGGAAGGGTGGGAGGAACATGGAATTTTCATTCTCTGGGCAC
CCCTTGAACAGTCTTACTATTAGGGCCCCAAATTTGTTCTAAGTGTGTGTGTGTGTGTGT
GTGTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAATTTTCTTTCTTCCTTTATATTCT
AAGTTCCTCAGGACAAAATTTTGGGTTTCTTTGTATTCTCCCTGCAGCTCCTCATGTAGT
[T,C]
CTAAGCAAATAAAGGAATTCATTAGGTCCTTGATTTCAGAAGCCTCCCAGTTCTCTATGT
AGGAGGAATCTTAGGGTGGCAAGATAAGTTGAGGGACTTTTCTTCAAGCACATTTCACAA
GTAAGAGAAAATGTTGACTGTGTATATCTAAGAATGGGTGGGGCTCAATGATGCCCCCCT
AAGTTACTCTTTACTATTATTGATTGATTGATTGATTGATTGAAGAAGCAATGTTTTGAT
TGATTGAAGAAGTAATGTTTCCAATGGCTACAGCAGACTGGAGCAAAAGAACAAAATGAA

48604 TTTTGGCTCTATCCTGGCTTCTTCACACAGGGGTGTCCAGTCATCTCATCCTGGTGGGAC
AGGGATAGAGCTGTGGCAGTGGAGATGAGGAAGCTCGCCTCCTAAGTGAGTCTGAATTCT
TAAATATGGAGCCACTCCATAATCATTTGGAGTGAATATTGGGCCATGGCCCTTTTTCTT
GCCAGCTGAGCTATGAAAAAAGGATGTCCTAAGACCAGAGGCTGTGGGACCATTCCCAGC
CCCTGCAGGAATCAAAGGAGCTGACAGAATTGTTTGTTTGTTTTTTTCACAAATTGAAAA
[-,A]
AAAAATGTAAAATTTTTGAAAAGAAAGCCTCATTGAAAAGAAATCCCTCTCCCCAGCTGG
GCTCCCAGGCAGCCTCCTGCAGAACATCCTTAGCATTGCAGAGTTGTTCCCATGGCAACC
GAGTAAGGGGCTTTTTGTTTTCCTTAGAAGATTGAATCCTTTCAACCAGAAGGTAACCAC
TGGTTCTTCCCCACAATCCACACTCCAAACCCCCTACCCTTATTTGACTACATGACTAGT
TTTGCATTTATGGATTTTTTTATGCCTAATTGAAAAAGGCTAAATATACAGAAACTGAGG

49560 TGAGGGGTTATGAGACCATAGGCTCATTTTGGGGGGGGTCTAAAATGCAGTATTTTTTGA
ACTGATATGGGGAAAAAAAGACATTTCTGAATTGTTGTCATGTTGCAGATTCTGGGCCGT
TCCAGCATAAGCACCTTTCTTAGAGTACTTGGCTTTGTGAAGTAGTCCTTATCCCCTCCT
TCCACTATTTTACATCAAGTTAAAATAGAGGAAGATGCCTAGAAATGGCCGTATAGACAG
AGAAAACTGCACTAAAACTCCCTCCGTCATGCCTGACTCCTCTCTAGACTATGACCATCG
[A,T]
GGGGCCAGAAATCATATCTTAAAGATCACTGTGCCTCCAGTACCCAGCACGGTGTTTAAT
AAATGTTTGTTGAATGAACGAACTAGTAAAATTTTCAAATCATTAGAGCTGAAGTATCCT
TTAAGATTCTTTAGTCCCTCATTTTACAGATAAGGAAGCTAAGGCTCAAGACATTGTGTG
GCTTGGCCAAAGGCACACAGCAAGCTAAAGGCAGAGGGAGGACAGGACCCGGCTGTCTCA
ACCCCCTGGCTGCTACACTTCCTGCAGCATTTCTAATTCTTTTACCATTCTTGCGAGGGA

52729 CCAATGGGGAAGCACCAGGGTCAGCCGCAAGGCAGAAGGAGCAAGAGGAAAACATGGACA
AGAGGCTCTACTGTGGATTCAGTGGCAAAGAATGGGAGGGGCAGAGTAAGCAGGTTTAGG
ATTATCGGGTTTGAATGACTTGATTGAGCTGTAGGGTGTAGAGACTGCCTCTACTGTCTG
GCACCAGGGGTAATTAGGGCAGCTGGATAGTGGTCTGGAGTGTGAGAGCTCCCTAAAGGA
GGTGGTTGGAGGTGTAGGTTTTGGATTGGTTGATCTGTATATGAAAGGTGCACGTGCAGG
[G,T]
TGAGTCCTCTACTATCACTAGAAATTGGCTGGTCCCAGGAGAAGTAGTCTCTCTAGAGAC
AGCAATGCCCCAGATGTCAAAGCATCAGAAAATACAGAAAAAAAATTAAAAGCATGATTA
ATTCATACTCACAGGTCTAGTTTTTGTGTAGTTAAGAGCAACCTAAAGAAGTTGATAACT
CGTGTTGCAGGTCAGGTTTCCCAGAAATCATATTCTCAGATGAAGATTTGCATGAAGGAG
GTTTAATGCTCAAACTAAGCCCTAAGGCTCCATACCTGTGGAGGAAGTGAAAGAAGCCCA

55031 TAGTGAGGCACACTTACTTCTTAATTTGTGCCACCCACTTTTCAGGCTCCCTTAGGACAG
CCTCCACCTGCTCCTACTGTGCTTCCCATCGTCCCTCTCCTCAGGCACAGGCTGAGGAGT
AATAAGAGCACCTGATATGTGTCAGGCCTTACTGTGTGCTAGGAATTGTGCTAAGTACTT
CCTATGAATTTTCCATTTATTCTTTATAATAACTTTGTAAAGTTAGAGCCATTATTCCAG
AAGGGAAAACCGAGGCAATGGGAGTCAAAGCAAAGAATTTGGGCTTTTAACCATTACACT
[A,G]
TTTTGCACAAGTAGCCAGTAATGAAAAGGCTGCTATCCGGAATCATCTTTGCAAAAGGTA
ATTTCTTTAGCACTTTATCAGAAGAAGGGGGCTCCTTCCTCAAATTCTGAGGGAAGAGAA
GTGGGGAAGAAAAGATGACTGAATCCAAAGCTCGGGCAGGGAAAGCACATCGAGTGCCAA
GTGCGCTGCGCTGGGGTCTAGTCCTGACTCAGCCGCCATCTTCCCAAGTGCTTCCTGGAA
TTCTCTCCTCTCGTGGGGCCTCAGCTCCTTCATCTTAGGAAAGAAGGGTAAAGATCTACA

55066 CACTTTTCAGGCTCCCTTAGGACAGCCTCCACCTGCTCCTACTGTGCTTCCCATCGTCCC
TCTCCTCAGGCACAGGCTGAGGAGTAATAAGAGCACCTGATATGTGTCAGGCCTTACTGT
GTGCTAGGAATTGTGCTAAGTACTTCCTATGAATTTTCCATTTATTCTTTATAATAACTT
TGTAAAGTTAGAGCCATTATTCCAGAAGGGAAAACCGAGGCAATGGGAGTCAAAGCAAAG
AATTTGGGCTTTTAACCATTACACTATTTTGCACAAGTAGCCAGTAATGAAAAGGCTGCT
[A,C]
TCCGGAATCATCTTTGCAAAAGGTAATTTCTTTAGCACTTTATCAGAAGAAGGGGGCTCC
TTCCTCAAATTCTGAGGGAAGAGAAGTGGGGAAGAAAAGATGACTGAATCCAAAGCTCGG
GCAGGGAAAGCACATCGAGTGCCAAGTGCGCTGCGCTGGGGTCTAGTCCTGACTCAGCCG
CCATCTTCCCAAGTGCTTCCTGGAATTCTCTCCTCTCGTGGGGCCTCAGCTCCTTCATCT
TAGGAAAGAAGGGTAAAGATCTACAGACAAATTGATCTTTAAGTATCCTTAGAGCACTAC

56912 TGAAATACTTTAAAACTTGTAGCTTCCTTCAGCACAGAAGTGGCTCTCTGAACCAATTTT
AAGCAATCCTGGCTCTATCTGTGCATGTTGATTTAGCCTGTGGTTATAGTGTTAACAATT
TAGTGATTCACCTCATTTTTAATCTCTCTTTCCCTTTAGCAGGATCATTTTCTCTGTGTT
AAGGGATCAACATTGAGGTAAGAATGGCTAAATAATAGCATCTTCTGGAATACAAATGAC
TTTATAAATAAAAGAAGATAAAAGGAAGAAGTAGGATGATTTCTCAGCTCTAATACACTT

[A,G]
GCAAATGCCATATGCTTTCTCCTGCGTGTACTGGTCAGGCCAGTTCTAGATACAATCATG
CGCTGCATAATGATGTTTTGGTCAACAGTGGATTGCATATGTGACGGTAGTCCTTTAAGA
TTATAATACCATATTTTTGCTGTGCCTTTTCTAGGTCTAGATATGTTTAGATACACACAT
ACTTACCATTGTGTTCCAATTGCCTACAGTTTCCAGTACAGTAACCTGTTGTACAGGTTT
GTAACCTAGGAGCAATAGGCTATACCATACAGCCTAGGTGTGTAGTAGGCTATACCACTT

58480 ACTGTCCTTCCTGTGTCTGAGGGAAGGCATGTAACTCTTGCTTATCTTCACCTGTGCTCT
AGATCCTGACCTTCTCTGGCAACCTCAGGGACCTTGCACCATCCATTCTTCTCGCCTAAT
GGCGAGACTCAGTCTCTCCCTCTCCCTTTCCACTCTCCCTTGCCATTCTTAGTATCTTTC
TACAAGCAGGTCTTCCAAAGTACTGCTTGAGGTCTGAGTTGGAGGGAACATGCCTCTACC
CTACTAAAAAGAGAAATTCCTCTGCAGAAGACCCAAGCTGACTGACAAATCCCTTTACTG
[C,T]
AACTGCAGCTCTAGCTCCCACCATTTTCCTGTACTTACTCTCCTGCTCAGGTTCCCTGGC
ATTGCTGATGTCTTTCAGCCTTTGTGCCCTGGCCCCTTTCCTCCTCTCCCCTCATCTAGC
ACTACCTGTCAAAATCAGGGACTTACTTTAAAATTTATCCCAAATTATCATTGCCATCAT
CTCCACTGTCACCTTATCATATGTTTGAATAGCGTTTCCATTTCCCAAATGTTTTCGCAT
GCACTTTCTCAATTGAGCCTTACGAATCCTAGAGCTGAGAAGGGTAACAATTTATGAGTC

61128 ATACAAAAATTAGCCAGTTGTGGTGGCATGCACTGTAGTCTCAGCTCCTTGGGAGGCTGA
GGCAGGAGAATTGCTTGAACATAGGAGGTGGAGGTTGCAGTGAACTGAGATTACGCCACT
GCACTCCAGACTGGGAAACAGAGTGAGACTCTGTTTTATATATATATATATATACACACA
CGTACATATACATGTATATATATACACATTATTATTGAAAGCAGCCAAAGAAAAATAACA
CATTATATATAGAGAAAGAGCAAATGATGAGTGACTTTATATGTATATATATGTGTGTGT
[G,A]
TATATATATAATGTGTATATATATACATATATATATATAGGTTAAGAACCTTCAGCACAT
GTATACCTATGTAACAAACCTGCATGTTCAGCACATGTATCCCAGAACTTAAAGTGAAAA
AAAAAAAAAAGAACCTTCTGCATGCCAGTAACTGTGCTAAGTGATTAGGATGCAATGGTA
ATAAAAACAAAGTCCCTCTCCTTAAAGAATTTTCTATTTAGAAGGGAAAACTGGTAAATA
AAAAATAAATATATAAATTACAATTTGTGAAAAGTGCTACACATGAAAGAGTGCTGAGAC

61320 TGTATATATATACACATTATTATTGAAAGCAGCCAAAGAAAAATAACACATTATATATAG
AGAAAGAGCAAATGATGAGTGACTTTATATGTATATATATGTGTGTGTATATATATAA
TGTGTATATATATACATATATATATATAGGTTAAGAACCTTCAGCACATGTATACCTATG
TAACAAACCTGCATGTTCAGCACATGTATCCCAGAACTTAAAGTGAAAAAAAAAAAAAAG
AACCTTCTGCATGCCAGTAACTGTGCTAAGTGATTAGGATGCAATGGTAATAAAAACAAA
[G,A]
TCCCTCTCCTTAAAGAATTTTCTATTTAGAAGGGAAAACTGGTAAATAAAAAATAAATAT
ATAAATTACAATTTGTGAAAAGTGCTACACATGAAAGAGTGCTGAGACAGACATCAATGG
ATAAACTTTAGATTGAGAAGGGCTCTGACAAAGCAACATTTAAGGTGCAACCTGAGAGAA
TAGAAGTTAAACAGGCAGATATTGGTGAAAGAGCAGTCTAGGCAGAGGGAACATCATTTG
CAAAGGCCCAGGGTAAAGAAGATCCTGGTAAGGAAATGACAGTGGAAGAAGGTTAGTGTA

61444 TATATATATACATATATATATATAGGTTAAGAACCTTCAGCACATGTATACCTATGTAAC
AAACCTGCATGTTCAGCACATGTATCCCAGAACTTAAAGTGAAAAAAAAAAAAAAGAACC
TTCTGCATGCCAGTAACTGTGCTAAGTGATTAGGATGCAATGGTAATAAAAACAAAGTCC
CTCTCCTTAAAGAATTTTCTATTTAGAAGGGAAAACTGGTAAATAAAAAATAAATATATA
AATTACAATTTGTGAAAAGTGCTACACATGAAAGAGTGCTGAGACAGACATCAATGGATA
[A,C]
ACTTTAGATTGAGAAGGGCTCTGACAAAGCAACATTTAAGGTGCAACCTGAGAGAATAGA
AGTTAAACAGGCAGATATTGGTGAAAGAGCAGTCTAGGCAGAGGGAACATCATTTGCAAA
GGCCCAGGGTAAAGAAGATCCTGGTAAGGAAATGACAGTGGAAGAAGGTTAGTGTAGCAG
GACTGTGGCTAGGGCGGAGAGGCAGGGAAGTAGTTTAGAATTTCAATGCAATAGGAAATA
TGGAAGATTGAAGGCAGTTTTGCATTATAAAATAATATGATTGCTATTTTAAAGCTACTT

62641 CCAGGCTGTACCAGGCACTCAGATATGACAGTGAATGAGATAGGCAACATCTTTGCCATT
GGAGAGCCTACACTGAAGTGGACATGAGGGAGTTGAAAGCAACTCTTATAGGAAATCATG
GTAAAGACGTCCAAGAGAAGAAAGATGAAGGGCAAACACATGCACGGATGCCAAACATCT
ATCAGAGAGAAAGGAATTTTCAGACCTGACCTGAATGATGAAAGGAGGTTTTTGGAAAGG
AAAATAGAAGGGAAGGACAAGGGAAATTATCTGGGCAGCAATATTTATCTGCTGTGGTGC
[T,C]
TCACTCTCTCTCTAATCCTTTTCCACCCCAGCCCCAAATTTGAAAGGATTGCAGGGAGCT
CCTGCTGGAGTCATTTCTGGTATTAAAAATGTACAGAAAGGAAAGCTTTGGTTCTGAGTT
TGCAGGCTTCCCTGTCTTTCATTCCTATTGTAGAAAGCAGCTTATATAAAAAGATGTGCT
GTGTGGCCCTTTGAGCTGCTGTGATTGTGTTAGGACCCCACTGGATGGTATTCGCATGAA
TTAATCTACTGTAGCATCTCTACAAATCAAGAGGCTGGCTTCTGTTTGAAATGTCCCAAG

63023 ATTAAAAATGTACAGAAAGGAAAGCTTTGGTTCTGAGTTTGCAGGCTTCCCTGTCTTTCA
TTCCTATTGTAGAAAGCAGCTTATATAAAAAGATGTGCTGTGTGGCCCTTTGAGCTGCTG
TGATTGTGTTAGGACCCCACTGGATGGTATTCGCATGAATTAATCTACTGTAGCATCTCT
ACAAATCAAGAGGCTGGCTTCTGTTTGAAATGTCCCAAGGCTTTGTGCACAGGGCAAGCT
AAATGTCTCCCTACAGTGAGACTGAAAATGCCTTGGGTGCCCTTGTCGATAGGATCTGAT
[A,G]
TATAGATGCATGTCTACAATTGCACAGTGGCTGCTGGCAACATTTATTACAATCTGAATG
TGAAATGGCTATTCTGTTCAAGGATTCTGATAAAAAGTATCAGCCACAGTAGATGTATAA
GGAGCCTGGTTTCACTGCAACTGACTACAGTTATCTGATTTTTTTTTTCTAGTTCATTTT
TAGTCTGTGGAGCAAACAGAGATTTCCTCCCCAAATGATGTCCTTTCTCAGTCACCAGGG

TGTGGTTATTTGGTTTTATGTAGAGGAGATAGAAACCAATCAGTCTAAATCATATTCTGT

63051 GGTTCTGAGTTTGCAGGCTTCCCTGTCTTTCATTCCTATTGTAGAAAGCAGCTTATATAA
AAAGATGTGCTGTGTGGCCCTTTGAGCTGCTGTGATTGTGTTAGGACCCCACTGGATGGT
ATTCGCATGAATTAATCTACTGTAGCATCTCTACAAATCAAGAGGCTGGCTTCTGTTTGA
AATGTCCCAAGGCTTTGTGCACAGGGCAAGCTAAATGTCTCCCTACAGTGAGACTGAAAA
TGCCTTGGGTGCCCTTGTCGATAGGATCTGATATATAGATGCATGTCTACAATTGCACAG
[T,C]
GGCTGCTGGCAACATTTATTACAATCTGAATGTGAAATGGCTATTCTGTTCAAGGATTCT
GATAAAAAGTATCAGCCACAGTAGATGTATAAGGAGCCTGGTTTCACTGCAACTGACTAC
AGTTATCTGATTTTTTTTTCTAGTTCATTTTTAGTCTGTGGAGCAAACAGAGATTTCCT
CCCCAAATGATGTCCTTTCTCAGTCACCAGGGTGTGGTTATTTGGTTTTATGTAGAGGAG
ATAGAAACCAATCAGTCTAAATCATATTCTGTTGAAATCAGAACCAAAGGATCCACAATC

64989 GGTTTAAGAAAGCAAGGTCAGGTAAGCTTCACAAAGTAAGTCAAGAAGTATTTTCACCTT
TATCTTCTGAAAGAATTTATGCAACGTTGAAATTATTTGTTTCAGAGATGGTCAACAGAA
TATACCAGAGAAACTATTTGGACTTAGAGCTTCCTTGGGGGAAGGTTTTTGATAAATAAT
GCAATTTCTTTAATACATAGTACTTATATTTTCTATCTTACCTTGTGACAATTCTGATGA
ATTGTGTTTTTCAAGAAGTTTGCCCATGTCATCTGAGTTGTTAAACTTACTACAACAAAG
[T,G]
CTTTGATAATATTCCTATATTAGCCTTTGAATGTCTATAAGATCTGTCCTGATGTTCCCT
CTCTCACTTTTTTAAAGAAGTCTTGCTAGAGGTTTACCAATTTTATTTTGTTTTATTTTA
TTTTATTTTTTCTTATTTGAGACAGAGTCTCGCTTTGTCGCCCAGGTTGGAGTGCAGTGG
CTCGATCTCGGCTCACTGCAAGCTCTGCCTCCCAGGTTCACGCCATTCTCCTGCCTCAGC
CTCCCGAGCAGCTGGGACTACAGGCACCAGCCACCATGCCCGGCTAATTTTTTGTATTTT

65929 TTAAACTTACACATTTCCCTTTAAGCACTGCCTTAGCTGTATCTCACAAATTTTGATATT
GTCTTTTCATTGTCTTTTATTCAATATATTCTAATTTTTCTTGTGATTTCTTCTTTGGCC
CATAGGCTGTTTAGAAATATGTAGTTAGTTTCCAAATATTCGAAGACTTTCACAGATACC
TTACTATTATTGATTTCTAATTTAATTCTGCTACAATCCAAGTATATACATTATAAAGTT
TCAGCCTTTTGAAATGTATTAAGAATATTACCAGAGATAAGAAGATAAGAATATTACCAG
[C,A]
GATAAGTAGGGATATTTCATAAATAATAGACGAATTGATTCATCAAGAATATACAACAAT
CATAAATGTGTATGTGTCTAATAACAGAGTCTCAAATTATATGAAACAAAACTGACAGAA
CTAAAGAGAGAAATGGCCAATCCCACAATCTTTATCTTTATCAGGTGATTTATCTTGGTG
AACATTCCTTGTGCTCTTGAAAAGAAAGTGTATTCTGTAGTCATTGGGTATAAAATTCTA
TATATGACAATGAGGTGATTGATAAAAATTATTTAGATTGTCTATATCCTAAGTTTTGTAG

66694 TCCGTTATTATTTATGCAATGTCCCTATTTATCTCTGGTCATATTCTTTATCTTGAAGTC
TTTTTAACTGATATGAATGTAGCCACTTCATCCTTTTTATGCTTACCATTTGCATAGTTT
ATATTTTTCCATTATCTTATATTCACACTATTTATCCCTTTATACTTAAGTCCATGTCTT
GTAGACAGTATGCAGTTAATTGTGTCTTGATTATTTTTACTCCTTTCTGACAATTTCTGC
CTTTCCATATAATATGCTTATCAATACAGTTGGAGTTAAATCTACCGTCTTGTTATTTGT
[C,G]
ACATCTCCCATCTTTTGTTGTTGTTCCTCATTTCCTTGTTTATTACCTTCTTTTCAGTTA
TTTTTTTTTTGTATTCCATTTTAATTCCTCAATTGGCTTTATAGCTATATATCTTTGTAT
TATTTTTTATTGTTTGCTCTAGGGATAGCAATATGTATACTTACCACAGACAATTTAGAA
ATCATATTGTACCACTTCACATAAAATAGAAGAAGCTTGCAGCAGTCTATGTCCCTTTAC
ACTCCCATTCTTTGTGCTATTGTTTCCGTATGTATTACATCACGTACATTGTAAAATCCA

66755 TTTTAACTGATATGAATGTAGCCACTTCATCCTTTTTATGCTTACCATTTGCATAGTTTA
TATTTTTCCATTATCTTATATTCACACTATTTATCCCTTTATACTTAAGTCCATGTCTTG
TAGACAGTATGCAGTTAATTGTGTCTTGATTATTTTTACTCCTTTCTGACAATTTCTGCC
TTTCCATATAATATGCTTATCAATACAGTTGGAGTTAAATCTACCGTCTTGTTATTTGTC
ACATCTCCCATCTTTTGTTGTTGTTCCTCATTTCCTTGTTTATTACCTTCTTTTCAGTTA
[T,A]
TTTTTTTTTGTATTCCATTTTAATTCCTCAATTGGCTTTATAGCTATATATCTTTGTATT
ATTTTTTATTGTTTGCTCTAGGGATAGCAATATGTATACTTACCACAGACAATTTAGAAA
TCATATTGTACCACTTCACATAAAATAGAAGAAGCTTGCAGCAGTCTATGTCCCTTTACA
CTCCCATTCTTTGTGCTATTGTTTCCGTATGTATTACATCACGTACATTGTAAAATCCAC
AATAGAGTGTTATAATCTTTTTCCAAATCCTTGTGTGAATTAAAAATTTTATGAGTAGAA

66879 CAGTATGCAGTTAATTGTGTCTTGATTATTTTTACTCCTTTCTGACAATTTCTGCCTTTC
CATATAATATGCTTATCAATACAGTTGGAGTTAAATCTACCGTCTTGTTATTTGTCACAT
CTCCCATCTTTTGTTGTTGTTCCTCATTTCCTTGTTTATTACCTTCTTTTCAGTTATTTT
TTTTTTGTATTCCATTTTAATTCCTCAATTGGCTTTATAGCTATATATCTTTGTATTATT
TTTTATTGTTTGCTCTAGGGATAGCAATATGTATACTTACCACAGACAATTTAGAAATCA
[T,C]
ATTGTACCACTTCACATAAAATAGAAGAAGCTTGCAGCAGTCTATGTCCCTTTACACTCC
CATTCTTTGTGCTATTGTTTCCGTATGTATTACATCACGTACATTGTAAAATCCACAATA
GAGTGTTATAATCTTTTTCCAAATCCTTGTGTGAATTAAAAATTTTATGAGTAGAAAAAT
ACATATAACATTTTATTCTTACCTACATACTTACCAGTTCTGCTTTCTTTTCATTCTTAC
CTGTTTCAGTCTTATCTGTAAACCCGTTTTCATTTGGTGTCATTTCCATTAGCATTTCAG

69156 GGCATGTCACTTCCTGCCCCAAACCCCTTGGTAGCTTTCCATTGCTCTTAGAATAACTTT
GTGATCTACAACATCTTCTTCAAGGCCCCGCATGATACAAATTCTGGCTATTTCTCTAGT
TTCTTATTGCACCACCTTGTCCCTCATCCACCTTTTTTTTAGTCTTCTCTCTTTCTTTGA

FIGURE 3, page 43 of 57

ACTTCTACCACCAGGTTTTTCACACGTTCTTCTTTCCCCATTAACAATGATCCACCATT
CTCTTTCTTTATCCACTGTTACTCATCCTCATAACTGAAACATCATTTCCTAAGGATGGC
[C,T]
ATTCCTGGTTCAGTCAGTCTATATTTCATCCCCCATCACATACTCTTGTTTTACCCTATA
TTTTTCCTTCAAAGCACTTATTTAAGTTGTAATTATGTGTTGTTTATTTTATGTCTGTCT
GCCCTCACAGAATCCACAGTCCAGGAGAACAGAAATCCTGCCTCTTTTATTTATACCACA
TCCACAGTATTATTAGTGCCTGTCACCTAGTAGGTATGCAGTATGTACCTATTGAATAAA
TGAATTGACTTCTGTCTTTTAGATCGTCTACTCATTTTATCATTGATGACAAACATAATA

69280 TATTGCACCACCTTGTCCCTCATCCACCTTTTTTTTAGTCTTCTCTCTTTCTTTGAACTT
CTACCACCAGGTTTTTCACACGTTCTTCTTTCCCCATTAACAATGATCCACCATTCTCT
TTCTTTATCCACTGTTACTCATCCTCATAACTGAAACATCATTTCCTAAGGATGGCCATT
CCTGGTTCAGTCAGTCTATATTTCATCCCCCATCACATACTCTTGTTTTACCCTATATTT
TTCCTTCAAAGCACTTATTTAAGTTGTAATTATGTGTTGTTTATTTTATGTCTGTCTGCC
[C,T]
TCACAGAATCCACAGTCCAGGAGAACAGAAATCCTGCCTCTTTTATTTATACCACATCCA
CAGTATTATTAGTGCCTGTCACCTAGTAGGTATGCAGTATGTACCTATTGAATAAATGAA
TTGACTTCTGTCTTTTAGATCGTCTACTCATTTTATCATTGATGACAAACATAATACCTT
ACATTCGTGTAGTCTTTTTCACTCCTCAAAGAGGATTTTCTGCATAGCTCCTCTGAGCCT
CACAAAACCCTTTAAGGAAGATTGTGAATATTATCAGATAAAGATTGTGAGACACAGAAA

70647 TCCAGTCATTTATAAAAGATGAAGAGGAGAACAAGGTAGGCCAAAGTGGCTTTGTACTAT
TAAAGGCTGCTTGATTTCTAAGTACATGTTCTTTGCCACCTTTCTGCCATTCCACATTCT
AGAAGCCATGGGTAAGTCAGCACAGGGATCTTAACATGATAACATTGGTTTTAGGAGGTC
TCGTGCATAATGGACCAGACTTAGAGCACAATGCTGTAAGGTAGTGATTTAGGTGAGCAG
CAGATTCTGGCTTTAGGAGTTTATTATCAGATGCTTTTTAAACGACTTGTGGCCCAGGAT
[C,T]
CCTGCACCCATGGGAAGCATTGTAGCCTTAGAACTCTGGGAATTCTGAATATAATTCCTG
AATCAATCGTAAGGATGCATATCTGATGCTTAGTGCAAACCAAGAGGCAGAATATTTGCA
GGCAGTGTATCCTTGAAAAACAAATCTAGGTCATTTTCCTGCCATGCTTCAAGCTTACTT
TTCCATCCTTCCTGATGGTAGTACTAACTACATTTGTAGACCATTTACGTGGTCAACACT
GTGCTAAGCTGTTAGCTTCATTCTCTATGAGACAGGCACTCTTAGCCCAACTTTACAATT

71867 TCTGTCTGGCTTTTCTCAACCTTTCTCTTCTGCACTTTCTTGGATATAATCAAAGCACTA
CCAGGAACTCCAGAGTCGGCACCTTTTCATTTTTGTGTTTTCATTTAATTATTTCTCAGC
TGCTAAGTGTTTGACTGTTTAAGGGACTCTAGTGGTAAATATTTGTCTTTAGCCTGGCAG
AAGCTGTGGTTTCCTTTGATGAGCTCACACGGTGTGGCTTTTAAGATGCTGCTGACCAGG
ACAGCTGACTGTCCCCAGTGGGTGCAGTCCCCAGCAGTGGGCTGGACCCCTTCCAGAAAG
[C,T]
GCTGCTGGGCCAAGAGGCTTCCTCCAACTTCCCGCTGCCCCCATCTAACCAACACCTCAG
TCTCTTCTCCACCTGCTTCCCTGCCCTCTTCCTTTCCCTCGCAGACACTTTCTTCTGCCT
GGCAAAAGGAATCTTGTTTCCATGGAAGCCTCATTAAATCTGCATCTTGCTCAGTTTGGG
TTTGATCACGGCTGCCAGAAGTATTTTTAGCCCATGCAGTTGCGTAATGAGATAGAGATT
GGGGAAAGGGGGAGGTGACTGTATAGGCAGAGGGTTTTTTTAAAAAAAAGTGAGAAAGAG

71900 ACTTTCTTGGATATAATCAAAGCACTACCAGGAACTCCAGAGTCGGCACCTTTTCATTTT
TGTGTTTTCATTTAATTATTTCTCAGCTGCTAAGTGTTTGACTGTTTAAGGGACTCTAGT
GGTAAATATTTGTCTTTAGCCTGGCAGAAGCTGTGGTTTCCTTTGATGAGCTCACACGGT
GTGGCTTTTAAGATGCTGCTGACCAGGACAGCTGACTGTCCCCAGTGGGTGCAGTCCCCA
GCAGTGGGCTGGACCCCTTCCAGAAAGCGCTGCTGGGCCAAGAGGCTTCCTCCAACTTCC
[C,T]
GCTGCCCCCATCTAACCAACACCTCAGTCTCTTCTCCACCTGCTTCCCTGCCCTCTTCCT
TTCCCTCGCAGACACTTTCTTCTGCCTGGCAAAAGGAATCTTGTTTCCATGGAAGCCTCA
TTAAATCTGCATCTTGCTCAGTTTGGGTTTGATCACGGCTGCCAGAAGTATTTTTAGCCC
ATGCAGTTGCGTAATGAGATAGAGATTGGGGAAAGGGGGAGGTGACTGTATAGGCAGAGG
GTTTTTTTAAAAAAAAGTGAGAAAGAGAAGGAAAACCTCTAAAGAAAAGAGTTTTATGGA

71901 CTTTCTTGGATATAATCAAAGCACTACCAGGAACTCCAGAGTCGGCACCTTTTCATTTTT
GTGTTTTCATTTAATTATTTCTCAGCTGCTAAGTGTTTGACTGTTTAAGGGACTCTAGTG
GTAAATATTTGTCTTTAGCCTGGCAGAAGCTGTGGTTTCCTTTGATGAGCTCACACGGTG
TGGCTTTTAAGATGCTGCTGACCAGGACAGCTGACTGTCCCCAGTGGGTGCAGTCCCCAG
CAGTGGGCTGGACCCCTTCCAGAAAGCGCTGCTGGGCCAAGAGGCTTCCTCCAACTTCCC
[G,A]
CTGCCCCCATCTAACCAACACCTCAGTCTCTTCTCCACCTGCTTCCCTGCCCTCTTCCTT
TCCCTCGCAGACACTTTCTTCTGCCTGGCAAAAGGAATCTTGTTTCCATGGAAGCCTCAT
TAAATCTGCATCTTGCTCAGTTTGGGTTTGATCACGGCTGCCAGAAGTATTTTTAGCCCA
TGCAGTTGCGTAATGAGATAGAGATTGGGGAAAGGGGGAGGTGACTGTATAGGCAGAGGG
TTTTTTTAAAAAAAAGTGAGAAAGAGAAGGAAAACCTCTAAAGAAAAGAGTTTTATGGAA

72369 TATTTTTAGCCCATGCAGTTGCGTAATGAGATAGAGATTGGGGAAAGGGGGAGGTGACTG
TATAGGCAGAGGGTTTTTTTAAAAAAAAGTGAGAAAGAGAAGGAAAACCTCTAAAGAAAA
GAGTTTTATGGAATTGGAAGAAGGATGGAGCACCTCTTTTGGGAGCATGAGGCTGGTGTT
CTCTGGTTAGCTCTTCCCACTGGAAGCCCATGGACACTTGCCATAATACCTGTCCTGGTC
ACATGTCAGGGGAACCTCTGATCTCCCTTTCCATGAGCTTAGTTGGCCCAGCCAGGGTGA
[C,T]
ACTTATGCTAGGGAGTGTGATTGATGTTGCTGCTTACAGATTTCCCCTCCCACAGACCTG
ATGGGGCAGCCAGGATAGTGGCAGAGAAGAAGACAGAGCAATAGCAGGAAAGAGAGGACA

ACACTAACACATTGGAGGTTTATGTTCAAAGACGGGATCTAGGGGGTCAGAGAAAGCACA
CCTACCATGTAATTGGTGCTGGAATCTGATGCCAAGTGCACCCTTGGCTTCTGAGGTTCT
GAGAACTCTTGCTTGTGCTTTTCAGCCAGACTATGCCCTCACCTGCCCCTGTACTTTAAA

72992 TGTTTGCATTGGATTGTTGGAGTGTGTGCATGTTGTTGTGTTCTTGTATTACAAGACA
AAGAGATTAAAAAAAAACCACATGCAGCTGTCACAGCTAATGTTTATTGAACTTTTACTA
TGCCACATGGTGTTTTAAGCATTCTATATGTGTTAACTCATTTTCCCTAATTCTATGGAC
TAGACACTTAAACAGTCTCCATTGTACAAACAAGGAAACTGAGGCACAGAGAGGTTGGGA
AACTCATTTGAGGTCCTCCAGCTAATTAATAGTGGAGCCAGGTTTTGTACCCAGACAACC
[T,G]
GATTTGAGAATCTGCAGTCCTAGATTAGTAACGTGTTGTTGGCCTGTCACACATTTTAAA
TGACATTCTGTACACAGAACCATTTATAGTAACTTTGTATTGTTGAGCTGAAAGCAGTCT
GCAGATGTGCTGCTGGGATTTCATTCATCTTCAAAGAGGTGTTTTTTTTTTTTTTTAAAG
GAAAATGCTTTTCTGAGGGTGGTATCTAAATTCATAAAAATCTTTACGATCAAGATTTTC
ACAAATTTCATTCTGACTCTGTTGCATTGCCCTTCTTCCCATATTCCCAGTTAGTTTGTA

73154 TTCCCTAATTCTATGGACTAGACACTTAAACAGTCTCCATTGTACAAACAAGGAAACTGA
GGCACAGAGAGGTTGGGAAACTCATTTGAGGTCCTCCAGCTAATTAATAGTGGAGCCAGG
TTTTGTACCCAGACAACCTGATTTGAGAATCTGCAGTCCTAGATTAGTAACGTGTTGTTG
GCCTGTCACACATTTTAAATGACATTCTGTACACAGAACCATTTATAGTAACTTTGTATT
GTTGAGCTGAAAGCAGTCTGCAGATGTGCTGCTGGGATTTCATTCATCTTCAAAGAGGTG
[-,T]
TTTTTTTTTTTTTTAAAGGAAAATGCTTTTCTGAGGGTGGTATCTAAATTCATAAAAATC
TTTACGATCAAGATTTTCACAAATTTCATTCTGACTCTGTTGCATTGCCCTTCTTCCCAT
ATTCCCAGTTAGTTTGTATTGATTGCTGCATCTCCCTTGAGCCCATGGTCCCCCACAACA
TTTCTTGCAGAACTGTGTCCTGCCTTCACACTGTCAGGCAGCAGGAGCCTCTCTAGCGGC
CAGCCCACAGTCCTGCAGCTCCTTCCTCAGGACGTTTAATTTCCCACATTTCTATGCAGT

73164 CTATGGACTAGACACTTAAACAGTCTCCATTGTACAAACAAGGAAACTGAGGCACAGAGA
GGTTGGGAAACTCATTTGAGGTCCTCCAGCTAATTAATAGTGGAGCCAGGTTTTGTACCC
AGACAACCTGATTTGAGAATCTGCAGTCCTAGATTAGTAACGTGTTGTTGGCCTGTCACA
CATTTTAAATGACATTCTGTACACAGAACCATTTATAGTAACTTTGTATTGTTGAGCTGA
AAGCAGTCTGCAGATGTGCTGCTGGGATTTCATTCATCTTCAAAGAGGTGTTTTTTTTTT
[-,T]
TTTTAAAGGAAAATGCTTTTCTGAGGGTGGTATCTAAATTCATAAAAATCTTTACGATCA
AGATTTTCACAAATTTCATTCTGACTCTGTTGCATTGCCCTTCTTCCCATATTCCCAGTT
AGTTTGTATTGATTGCTGCATCTCCCTTGAGCCCATGGTCCCCCACAACATTTCTTGCAG
AACTGTGTCCTGCCTTCACACTGTCAGGCAGCAGGAGCCTCTCTAGCGGCCAGCCCACAG
TCCTGCAGCTCCTTCCTCAGGACGTTTAATTTCCCACATTTCTATGCAGTTACCTCACAG

74149 TTTGCTCAAGGTCACATAACTAGTAAGTGGGTGGAGCTGTGATGTGAAACTGGGCAGTCT
GATTCTGGGACCTGTGCTCTTAATCACCAATCTATATTGCCTCCTACTTGAAAACATCCA
GGGAAAATGTTGAGATAGATCAGCTGAAATCTTCTTGCACAGTAAAGCAGGGGCCACCTG
TCCTGGAGTTACATTCATCTTGTTCATTGTCAACGATTTGTGTTCAGTGACACCCTCTTC
AGCCCAAGAACTTACCTGGGTGCTGTGACAATTGGACATGACTAGGAACAACCAGTGACA
[T,A]
TGTAGCCCATCCAAACACAGGGTAGGAAGTGGATGCTTGTCACTCTCTTTTGGTTATAAG
AAGCAGGAACCCAGTAAAGGCACCTTTTATATATCTATAAAGTTGAATATATAAGATATA
TGGGGGCCAGGCACAGTGGCTCACACCTGTAATCCGAACATTTTGGGAGCCCAAAGCAGG
TGGATCACCTGAGGTCAGGAGTTCAAGACCAGCCTGACCAACATGGTGAAACCCCATCTT
TACTAAAAATACAAAAATTAGCTGGGCGTGGTGGCACACACCTGTAGTCCCAGCTACTTG

74171 GTAAGTGGGTGGAGCTGTGATGTGAAACTGGGCAGTCTGATTCTGGGACCTGTGCTCTTA
ATCACCAATCTATATTGCCTCCTACTTGAAAACATCCAGGGAAAATGTTGAGATAGATCA
GCTGAAATCTTCTTGCACAGTAAAGCAGGGGCCACCTGTCCTGGAGTTACATTCATCTTG
TTCATTGTCAACGATTTGTGTTCAGTGACACCCTCTTCAGCCCAAGAACTTACCTGGGTG
CTGTGACAATTGGACATGACTAGGAACAACCAGTGACATTGTAGCCCATCCAAACACAGG
[G,A]
TAGGAAGTGGATGCTTGTCACTCTCTTTTGGTTATAAGAAGCAGGAACCCAGTAAAGGCA
CCTTTTATATATCTATAAAGTTGAATATATAAGATATATGGGGGCCAGGCACAGTGGCTC
ACACCTGTAATCCGAACATTTTGGGAGCCCAAAGCAGGTGGATCACCTGAGGTCAGGAGT
TCAAGACCAGCCTGACCAACATGGTGAAACCCCATCTTTACTAAAAATACAAAAATTAGC
TGGGCGTGGTGGCACACACCTGTAGTCCCAGCTACTTGGGAGGCTGAGGCAGGATACTTG

74918 TAACAGGTGCTGAAAACAGGAACTGGGAAGTTGCCAGTACCTTCCTGTCTTTTCCCCTGG
AACCAAACGGTTTCTTACTTGCTTCTCTCTGCACCTCTGTCTCATTTCCCTCTCTCTTCA
GATGATTTTTCATTGTTGCATCACACACATAGAAAAATCAGGATCCACCCTCCCAAGTTT
ACATATCGTTGTTTCAGGCAGCCATAGTATCCTTAAAACTCCACATTCCAGGGAGAAAGC
TTGGGTCAAGGATTCAGCCAAAGGGCAGCGAAATGGAGTAAAGATGCAACTGCCAGGTCT
[A,G]
TGGGCAGCAAGGAGGCCGGGAAGGAAGCCGCTGTTGTGGTCCAAGTGACAATTCAACAGC
TCAAAGCATAAGTAAGTTGTGTGCTTTTCACAGATGGAGAAACTGAGGCACAGAAGGAAC
CTGGCTGGGGTCCAGGTCTCTGGCCTTTGTGTCAATGCTAGGTCACTGGATGTGGCGTCT
GATTTCTACAGGAAATGTGGTTTCTCTACTTTGTCCCAGAGCCCACTCAGAGCACTGGCT
GGCCAGGGGGTCCTAGGGCCCTCTTAGGATAGTCTCAGGCCAACAGCCCCAGGACAGAAG

75386 GGATGTGGCGTCTGATTTCTACAGGAAATGTGGTTTCTCTACTTTGTCCCAGAGCCCACT

FIGURE 3, page 45 of 57

CAGAGCACTGGCTGGCCAGGGGGTCCTAGGGCCCTCTTAGGATAGTCTCAGGCCAACAGC
CCCAGGACAGAAGCAACCAAAGTGAAGTTATGAAAGAAAGCTCTTTGCTGATCTGTCAAT
GGCACCCTTGTAGAGCCAATACTTAGAACACCTGGATTTGAATACTCATCTCCAAAACCT
GTGTTCTTTCTACCACGTGACAAGCCCTTGTAAACCTCACAACGTCTCTATGAGGTGAGC
[G,A]
CTTGCAGATCCACACTTTAGATAAGCAAATGGAGGCTCAGAGGGTAAGCAGCTAGTTCAA
GGTTATGCACCTGAGCCAGGATGTGGACACAGCTCTGTGTCTGATTCCTAAGGGCCTGTG
CTTTAGCCACTTTGCAATACTGCTGCTGTCTGCTTCATTTCCTCATCTGTCAGATGGGAA
CGATAATACTCAACTCACATGGATACTGTATGAGGAAAAACAGATAAAAGAAGAGAAAGT
GCTTTGAAAACATAAGCAGCCCTGGCAGATGGGAATTATTTTTGCTGCTGACACACATCC

77751 CAATGCCAATGTTTCTGAAGCCCATATTAAATGCCAAAATCTGAGTCAGCTACTGGAGGT
AGAGACATGAATAAGATGGTCCATATTATTTTAGAGGATTCTTTGGTTGCAAAGGGCAGA
CACCCAGCTTGAATTCACTTTGGAGAAATTGGGATTTTTTTGGCTTGCATAAGCAAAGCA
TGAGAAAGAAAGTTCCAGGGATGATGAAAACCAGGAATGCAAATGTCTCCAGAATTCTTT
CTTTTTTCTTTTAGGCCATCTTTTTTCTCTCAAACTGGTTCCCTCCACTGGGCTGGAGAC
[G,A]
TTACTACCAGCAGCACTCAGACCCACATCTTCAGTTTAAATGTTGGAAATGGACTGTCAG
AGAACATTTAGGCCATTCATTCTGTGGGAGAGATAGGCTATGTAAAAAGATAGCCACTCC
CATGTGAACAATGTGGTTAGGATTAGAGGCATGAATATACCCCAAACCAGGGGTGTGGGA
AGGAGGTTGACACTCTAGGTGATAATACCCAGACCTTAAGGAGCTTTCTGTCTAGAGGGA
GGTATGGACATGGACAAGTAATCAACAGCTACAAAGCAGAGCTGCCAGCTCTGCAACACA

78264 ACCTTAAGGAGCTTTCTGTCTAGAGGGAGGTATGGACATGGACAAGTAATCAACAGCTAC
AAAGCAGAGCTGCCAGCTCTGCAACACAAGAGCCCTGAGAGGCATGACAGGGGCAGGGTG
GGGATCCATGTGGGTCTGGATTGAAGTGAGGAGGGGCATCAGGAAAGCATTCCAGGAGAG
CTGAGGGACACTTGAGCACACCCTCAAAGAATGACTGGGGTCATGAGGTATACAAGGGA
GGAAGTGCACCCGAGACAGAAACAATCACATAAGCAAAAATGCAGAAGAATATGAGGATC
[G,T]
GGGAAGGGCAAGTAGCTCAGTAGTGTTGGAGGCCAAGGGACACGAAGGAAGGTGATAAAG
CCCTGATGTTAAGGATAGAAAAATCAAAGTCCTTTGAAAATCATGTGGAGTTAGGATCTC
AAGAACCCTACAAGGATTTCTTTAGAATAGAATCAAAGAAAAACAAAGTTTACAGTCTGT
GAGGGTTGCATAGGAAGTAACGTGGTGAGAAATGTTGGCTTGAGAACCACATATCCATAA
CACAATGGTGTTTTAGAGGATTTGGGGGAAGGGAGAGAAAATCTCAAATTGTCTCAGTAA

80986 GCATCATATTGCATGAAAACAGCAAACGGAAGTCACAATGGCTCGACGGTGTAATGAAGC
CACACAATATGTATTAAACACATCATCTACACAGATGGATTCAAAGATACCTTCTTTGTG
TCTAAGTCCCAAATCTGTGTTTCCTGGCTCTGTTCCCTCATATCTAGTCATTCTCCAAGT
CAGCATGCCCAACTTGAAAGTGTCATTTTCAAAACCTGCTTCTTCTCTTCTGGAAGTTCT
TCCTCTGCCCATTGCTCCACAATCCCCACCTCTTTCACCCAGTAGCAAACCTTAAATTTA
[T,A]
CTTTTACTTTGTCTTACTTCCCCTTCTTATATTCAAAATGTTTCTCACTTGCATCTCTTT
TCATTCATTTCATAAGCATTTATGAGCTCCTGTTATGGTTTGGAAACTGTTCTTCATGCT
GGAGGTGGTCTTATAAACAAGTAATTTCAATTGAGTATTTAGTATGTTAAGTGCCATCCC
AAAGGCAAACACCAGCTGTGGGAGGCTCCCCAAATCAGTCTAAGGAAGTTGGGAAAAGCA
TCTCAGAGAAGATGGTGTCTGAGATGGGGAGGATGTGTGGAACTGGGCAAGGAAGAGAAC

83609 TTTGGGCAATTGTAGCAATTTTAAAACTATGTTAGATGGCTAGAGATTCTTGAGAATATT
TCTTTTCTTGGAAAATCATAAGGCTTTGGATAGTGGTACCTATAGAAGCTGACATCAGCA
GCAGCCTGCCTCCAGTCGATCAGGGCCTTTGGAACTTCACGGGGCTCCTCTACTGACAGC
CCCATCGGTTTCCCTCCAGCACACGTAACTCAGCATTGACTCTGGGTAGTAGAGGGTGGT
TTATGGAATCTGATTCATCTCAGAAAGAGGTGGATGCAAACACATTCCCAGAGCAGAAGG
[C,T]
TTGGCATGTCTGGTCTTAGGCAGAGGGAACTGGAGATACTTGTCCTATTGTTCTTGAGAT
TCCAGCAAAAATAGCCCATTACAGAGGAAGAAGATATCAGGTCAAATGAAGGCTTTGGTG
CTACAACATTGTCTTAGAAAAAAAAAGAAAGAAATTGGCCAAGTGCAGTGGCTCAGCACT
TTGGGAGGCTGAGGGGGGCAGACCACTTGAGATCAGGAGTTCGAGACCAGCCTGGCCAAC
ATGGCGAAACTCCGTCTCTACCAAAAAGTATTAAAAAATAGCCGAGTGTGGTGGCGGGCT

85271 CCTTGGGGCATCACATTAAGTAGTTACCAGATTGAACTGCAAACATTGCTATCCAGGAGA
AATCAGGTCAATATTTCACCTTCATGGCAATACCAGTACAGTCCAAGGAGAATGCATAGA
AGGAAAGAAATCATAATCTGATTCTATGTGTTTTTTTAGTAGTAAATAATAATAATTATT
ACTATTCCTATACAATTTTGTGTGTTGGTGTGTTTTGTTTTGTTGTGCATGAAAAATGGG
GTGCTAATCTATTCCCCTTCCCAACACCAGTGCTCAGAAGAAATTTCCACAGATAGAGAA
[G,T]
CTATAGGTTATGAATTTGGCCTTGATGGATTCTGGGTCACTATTTCTCAATGTTTGTCCA
TGTCATGTGAAGCTCTTAAGATAAAGAACAATGTCTTACTCGTCTTTTTAACTTCTTTAC
CCCCTAATGCCTATCACATACTTTGCCCATGGAAACTCAATAGACATTTGTAAATGGAAT
TTAATTTCTGAGGTCCAGTAAAGCCTTTTTCCATCCTTCCCCTACTACACAGTTTGTCTA
ACCATGTCTTCCCTTCCATCATCCACCTTATAAACGTTATTACTCATTCTTCCATCACAT

87770 CTCCCTACCTGTCCCTCCTTGACCCCAGGAAAAATTGCCGGGATATGAAAGTTAATTATG
ACCCAAGGGAATTGGTACAGATGGGGAAGAAAGAAATGCATTCAAGAGCATTTCCATCAG
TATTGAAATTACACAGAAGGCTGGTGAATTTGGGCTATCCATTCTTGCCTCCCTCTGTGC
CCATAATTCCTTGGCCTCCTTCAATTTCATTTTCCCTTTGGTTCAGAGGAATGCTTGATG
GCTTAAGCTAGCCTCAGTTGGCCAAGCATTGGAGAAACAGAGAGGTGTATGACACAGCTA
[C,T]

ACTCCCATGGGGCTTACAGGGCAAGGTGAGAGAAGACAGAAGTTGTATGTGCTGGGTGCC
ACGTGGTAGCTACAAACTAGAAATGAGACCAGGTTCGGAAGAGGAAGAGGGCTTGCAGAC
CTGAGTCATGGGGACAGTTTCTTCAGGAAATGGGATCTCAGCTCTGCCTTGTATGCAGGG
CTTACATAATAAATATGTTTCATTGTTGTTGTTGTTATTGTTGATTTAATAAGATTTTGT
TTTAAGAAGATTTTGTAAAAACAACTGAACAAATGCAATCTCCTGCCAGAGCAGGCAGCA

87837 GGAATTGGTACAGATGGGGAAGAAAGAAATGCATTCAAGAGCATTTCCATCAGTATTGAA
ATTACACAGAAGGCTGGTGAATTTGGGCTATCCATTCTTGCCTCCCTCTGTGCCCATAAT
TCCTTGGCCTCCTTCAATTTCATTTTCCCTTTGGTTCAGAGGAATGCTTGATGGCTTAAG
CTAGCCTCAGTTGGCCAAGCATTGGAGAAACAGAGAGGTGTATGACACAGCTACACTCCC
ATGGGGCTTACAGGGCAAGGTGAGAGAAGACAGAAGTTGTATGTGCTGGGTGCCACGTGG
[T,C]
AGCTACAAACTAGAAATGAGACCAGGTTCGGAAGAGGAAGAGGGCTTGCAGACCTGAGTC
ATGGGGACAGTTTCTTCAGGAAATGGGATCTCAGCTCTGCCTTGTATGCAGGGCTTACAT
AATAAATATGTTTCATTGTTGTTGTTGTTATTGTTGATTTAATAAGATTTTGTTTTAAGA
AGATTTTGTAAAAACAACTGAACAAATGCAATCTCCTGCCAGAGCAGGCAGCAGCAAAGG
AGATTAGGAATATAACCCCCTTGGAGACGTTCCTTCACCTACCTGGTGCTGGATTACCTA

87866 TGCATTCAAGAGCATTTCCATCAGTATTGAAATTACACAGAAGGCTGGTGAATTTGGGCT
ATCCATTCTTGCCTCCCTCTGTGCCCATAATTCCTTGGCCTCCTTCAATTTCATTTTCCC
TTTGGTTCAGAGGAATGCTTGATGGCTTAAGCTAGCCTCAGTTGGCCAAGCATTGGAGAA
ACAGAGAGGTGTATGACACAGCTACACTCCCATGGGGCTTACAGGGCAAGGTGAGAGAAG
ACAGAAGTTGTATGTGCTGGGTGCCACGTGGTAGCTACAAACTAGAAATGAGACCAGGTT
[C,T]
GGAAGAGGAAGAGGGCTTGCAGACCTGAGTCATGGGGACAGTTTCTTCAGGAAATGGGAT
CTCAGCTCTGCCTTGTATGCAGGGCTTACATAATAAATATGTTTCATTGTTGTTGTTGTT
ATTGTTGATTTAATAAGATTTTGTTTTAAGAAGATTTTGTAAAAACAACTGAACAAATGC
AATCTCCTGCCAGAGCAGGCAGCAGCAAAGGAGATTAGGAATATAACCCCCTTGGAGACG
TTCCTTCACCTACCTGGTGCTGGATTACCTAAAAGCTTCAGCTAAGTAGGGTCACCCCCC

88238 CTTGTATGCAGGGCTTACATAATAAATATGTTTCATTGTTGTTGTTGTTATTGTTGATTT
AATAAGATTTTGTTTTAAGAAGATTTTGTAAAAACAACTGAACAAATGCAATCTCCTGCC
AGAGCAGGCAGCAGCAAAGGAGATTAGGAATATAACCCCCTTGGAGACGTTCCTTCACCT
ACCTGGTGCTGGATTACCTAAAAGCTTCAGCTAAGTAGGGTCACCCCCCCAAGAAATTAT
TTTAAAAAAATTGAAATCTGATATTTTTAGAAAATCTTATCAAGGATATTTAATTGGACT
[A,C]
TTTACACCTATTTAGGGTCAGTCGGTTTTGGACAAGTATGCAGGGGTCTTGGAATCAGAC
CACTGGGGTCAAATCCTAGTTCTGTCACTTCCTAGCTGGGTGACCTTGGACAAAGTTACC
TGACTTCTAATAGCTTCAGATTCCTCATGGGCAAAATAGAAATGCTACTAGTACTTAATA
GTGCTCTGAGAAGGATTCAATGAGAAGGATTAAATGTATGTAAAGCACAGTGTTTGCCCA
TAGGAAGCTGTTATTTATAAGGGAGGGGGAGCATCCTAAGGTCCTCCGAATTTAGGAGAAC

89219 AGAGGGCAGTTGGAAAACTTCACAAGACAATCCAGCCTGATTGTTTTGACATGCCTGACT
TCAGGCTGCTAAAAATGAGCTCGAGGAATCAGATAGGAAAAAGAGATAGGTGATGCAATT
TTATTCCATCTCCCAATTTTCTGAGTCAAGAGTTGTTTGTTTAACTCCAGTTAAATTAGT
ATTTATCCAAATTTCCTGGGTGCTTGTCCAAAGAAAAGTACCCCAGATCTACAAATTAGA
ATCTGGGACTGGGACTTAGGAATTGGCACTTTTACAATTATACCAGATGTTTCTAATATG
[A,G]
GTACTTCAACCACTACCCTTATAGAAGTGCTGCCTAGGACCCTCTCTTCTGGCAGGTGAA
GTGGAAGGAGGTTTTGTCGAAGGGAGATTCTCCACTTCAACTTGAGTGTCTTGGCTTGTA
TCCGCTTTGTTTGGTTCTATTTCACCAAAGGCTTTCATCTTCACATAAATTTTCTTCAGC
TTTAAATAATTAGTTTTGGTAACCATTGGTATACTGGAAAGAACATTAGATTTGGAGTCC
AGGTGGCTTGAGTTCAATTCTCTGCTCTGCCATTTACCAGCTGTGTGACATTGGGCAAGT

89331 ATGCAATTTTATTCCATCTCCCAATTTTCTGAGTCAAGAGTTGTTTGTTTAACTCCAGTT
AAATTAGTATTTATCCAAATTTCCTGGGTGCTTGTCCAAAGAAAAGTACCCCAGATCTAC
AAATTAGAATCTGGGACTGGGACTTAGGAATTGGCACTTTTACAATTATACCAGATGTTT
CTAATATGAGTACTTCAACCACTACCCTTATAGAAGTGCTGCCTAGGACCCTCTCTTCTG
GCAGGTGAAGTGGAAGGAGGTTTTGTCGAAGGGAGATTCTCCACTTCAACTTGAGTGTCT
[T,C]
GGCTTGTATCCGCTTTGTTTGGTTCTATTTCACCAAAGGCTTTCATCTTCACATAAATTT
TCTTCAGCTTTAAATAATTAGTTTTGGTAACCATTGGTATACTGGAAAGAACATTAGATT
TGGAGTCCAGGTGGCTTGAGTTCAATTCTCTGCTCTGCCATTTACCAGCTGTGTGACATT
GGGCAAGTTGCCAACCTATCTATGTCATTTCCTCATGTAAAGATAATCCCACTTCACCAG
GCCACTTTTGAGGACCCAGTGAAATGATGTGTAACCATTTTAGGAACACTGGATCATTCT

90794 GATTTGGCCAAGGTCACACAGCTATAAGCAGTAGAACTAAGATTTTAACTCAAGTTTCTA
TGGCCCCAGAATTTATGTGTTTCTCTCTCCATACCACAGGGACAGGTGCAAGTGAGAGAT
TTTGCTGGAAGCACTGGGCTTTTGAGCAGGCCATATAAAAATTCTGAGCCCAGAGCTCA
ACTAAATTATTGGAAGAGACTGGGCCAAATATAAGGCTTCTATCTAAGCAGCACCTGTGT
TTCTCAAGGACTGAGGAAAATGAAGGGGGAGGGTTGGCAAGGCTGCATTTCCCAGGGTGC
[A,G]
TGATTATATGGCATGGGGGTGGGGGCCATTATGATGCCCGGACATGGAACTTACACCAGT
GCAGAAAGGGTGTGATTAGAAGCCCTAAGCCAGAGAATGTTCAGTGTGATAAATGCCATT
ATTTTTTCCCTCATTCATTCAATAGATTTTTTTTTAGATGGAGTCTCACTCTGTCGCCC
AGGCTGGAGTGCAGTGGCACCATCTCAGCTCACGGTAACCTCTGCCTCCTGGGTTCAAGC
AATTCTTGTGGTCCAGCTTCCTGAGTAGCTGGGATTACAGATGTGCACCACCACGCCTGG

92404 TGCTCTGCTACCTACCTGCCAGCTGTTTCCCAGGGATGTGGTAAAGATGAATGGGCAAGA
TCTGGGAAAGTGTTTTGAAATCCTTGATTAAAGGCCCTCCAGGCAGATGTAGAATTTTAA
ATGTGTTATATTACTGCCACTATTGTTATGCTTTCTTTTATCACCCCAGAATTTCACCAT
CTCCTGTTTCAGGTGAACGAGTCTGCCTGACTCTTACCTGCCCTGAATGGCATTGGAAAG
GTAGCAGCCCTGAGATGTGCCATATAAACAAACATGTTTTTAACCAAGGGATCAGGAGGC
[C,T]
TTCCTGGCTGGCTCCTGTCAGCTGGTCATCACCTCTCTATAACTCTAGGCTTTCCCAAGC
TTATTTTATTTCCATCAATAGGACAGGAATATGTAAATGTCCTGCTTGAAATGAGTATTG
GCTACAAGCCATCTGCCTCTGAACAGAGGTGAAAAGTGGAAATCGGAGGAAGGGCAGATG
TCTTTTGCAAGGGAAACAGACTGTTTTCTGCCACTGCACTCTGCCCAGGCAAAAGAGTAA
AGGAACAGCACTCAGGAGAATTCACTGAAGCGAGGGCAGGGTGCAAAAGGAACTTGAGAA

92672 CAAACATGTTTTTAACCAAGGGATCAGGAGGCCTTCCTGGCTGGCTCCTGTCAGCTGGTC
ATCACCTCTCTATAACTCTAGGCTTTCCCAAGCTTATTTTATTTCCATCAATAGGACAGG
AATATGTAAATGTCCTGCTTGAAATGAGTATTGGCTACAAGCCATCTGCCTCTGAACAGA
GGTGAAAAGTGGAAATCGGAGGAAGGGCAGATGTCTTTTGCAAGGGAAACAGACTGTTTT
CTGCCACTGCACTCTGCCCAGGCAAAAGAGTAAAGGAACAGCACTCAGGAGAATTCACTG
[A,C]
AGCGAGGGCAGGGTGCAAAAGGAACTTGAGAAATTGGTACTGGGACCCAAAATCAGATTC
TGGCATTTCTGGGAAAAGAAATGGGCATGGGTGGGGGTTTTATCTGTCAATAAAAGCATC
CAGAATGGGGCTAGAAGGAAGTAAATTCAGTTGCCACCTCTGCCTACTGGACAGCCACGG
AGAACTTCTCCTTATCCAAGGTCGAGGAGCCCTCCGGAGTACATACTGATACCATTGGTT
CTCCCACACATACCCCCATGGAGATAAAAACAGGACCCTGGAAGCCCTGTCCGTGTTTAA

92684 TAACCAAGGGATCAGGAGGCCTTCCTGGCTGGCTCCTGTCAGCTGGTCATCACCTCTCTA
TAACTCTAGGCTTTCCCAAGCTTATTTTATTTCCATCAATAGGACAGGAATATGTAAATG
TCCTGCTTGAAATGAGTATTGGCTACAAGCCATCTGCCTCTGAACAGAGGTGAAAAGTGG
AAATCGGAGGAAGGGCAGATGTCTTTTGCAAGGGAAACAGACTGTTTTCTGCCACTGCAC
TCTGCCCAGGCAAAAGAGTAAAGGAACAGCACTCAGGAGAATTCACTGAAGCGAGGGCAG
[A,G]
GTGCAAAAGGAACTTGAGAAATTGGTACTGGGACCCAAAATCAGATTCTGGCATTTCTGG
GAAAAGAAATGGGCATGGGTGGGGGTTTTATCTGTCAATAAAAGCATCCAGAATGGGGCT
AGAAGGAAGTAAATTCAGTTGCCACCTCTGCCTACTGGACAGCCACGGAGAACTTCTCCT
TATCCAAGGTCGAGGAGCCCTCCGGAGTACATACTGATACCATTGGTTCTCCCACACATA
CCCCCATGGAGATAAAAACAGGACCCTGGAAGCCCTGTCCGTGTTTAACCAATGGGATTG

93132 CTGCCTACTGGACAGCCACGGAGAACTTCTCCTTATCCAAGGTCGAGGAGCCCTCCGGAG
TACATACTGATACCATTGGTTCTCCCACACATACCCCCATGGAGATAAAAACAGGACCCT
GGAAGCCCTGTCCGTGTTTAACCAATGGGATTGAAACATGGAAATGAACTGCCCCACAAT
CCACCCTGTGAGAGACCAAAGAGCAGTGTTGGATTAACAGGGAATGTTACCCTGAAAAGG
CATTCAGCTTCCACTGGGGCAGCAGGTACAGTGCAAAGATGATCCCACTTAAATTCCTAA
[G,C]
ACAGGAAATAAGGAAAGATGTTGTGGAAACTCAAGACCTCTCAAAGCATACTCCTTTGTA
GTTCTTCCGCAGACCAGACCACGGAATTCAGAAAACACCCTACCTGGTTCCAAACCAGCA
CCTGCCAAACTTCTCACCCTCTTCTGACCCTGTCCTGGGAGTTAAGAAAAAAAAAAATCAC
TTTATTGGTTGCTCCAGTTATAACTTAAACAGACAGACCATCATCAAATTAAGTGACATG
TACGACTGCTTATTGTATGCCAGTTACTGTGCTGTGGGGTTTTGGTTCCATTATCTCATT

93537 TGGTTCCAAACCAGCACCTGCCAAACTTCTCACCCTCTTCTGACCCTGTCCTGGGAGTTA
AGAAAAAAAAAAATCACTTTATTGGTTGCTCCAGTTATAACTTAAACAGACAGACCATCAT
CAAATTAAGTGACATGTACGACTGCTTATTGTATGCCAGTTACTGTGCTGTGGGGTTTTG
GTTCCATTATCTCATTTAATCCTCTCAAAAACCCTGTTAGGTAGGTTTTATTATTGCACT
CATCTTAGATTAAGGAAACTGAGGCTCATAGAGATTCGGTAATTTGTCAAAAGCCCTAAA
[A,T]
CATAATTACTGCCTCCAGATGTCTCTGATTCTAAGGCCCAGGCTCTTAATCAGTAAATGA
TCAAATGAATAATGATTTTCATGGCATCTGTCATCGGAAAGAACAATGGAGAATATGCTT
AACCAAAGTCATAACCAAATAAATGAACTTGACAGCAGAGCCGTGATTCTAGCCAAGATG
ACTATTTTCATGCATGTTTTGAAGGCCAGGAAAAGGAGGTTAGACTTGTTTGGGAAGGGA
AACAGGAGCTATCAAGGTGAACTTTTCCTAAGAGTAGCCCAATAATAGTGCTCGGGAGGG

93557 CCAAACTTCTCACCCTCTTCTGACCCTGTCCTGGGAGTTAAGAAAAAAAAAAATCACTTTA
TTGGTTGCTCCAGTTATAACTTAAACAGACAGACCATCATCAAATTAAGTGACATGTACG
ACTGCTTATTGTATGCCAGTTACTGTGCTGTGGGGTTTTGGTTCCATTATCTCATTTAAT
CCTCTCAAAAACCCTGTTAGGTAGGTTTTATTATTGCACTCATCTTAGATTAAGGAAACT
GAGGCTCATAGAGATTCGGTAATTTGTCAAAAGCCCTAAAACATAATTACTGCCTCCAGA
[T,C]
GTCTCTGATTCTAAGGCCCAGGCTCTTAATCAGTAAATGATCAAATGAATAATGATTTTC
ATGGCATCTGTCATCGGAAAGAACAATGGAGAATATGCTTAACCAAAGTCATAACCAAAT
AAATGAACTTGACAGCAGAGCCGTGATTCTAGCCAAGATGACTATTTTCATGCATGTTTT
GAAGGCCAGGAAAAGGAGGTTAGACTTGTTTGGGAAGGGAAACAGGAGCTATCAAGGTGA
ACTTTTCCTAAGAGTAGCCCAATAATAGTGCTCGGGAGGGAGTAATGTGTGCAAGAATAG

95067 AGAGAAATGGAAGCAGGGAGATAAATTAGGTGGTTATTGCAAGAGGCCAGGTAAGAAGAG
AAAGTGGTTTAAGTAGGGTGGTGTGGCAGAGAAGACGGTTCCAAGCAGAGGGGGACCACG
CTGACAAATAAGCGCGGGCCACTCACGCAAGCCCAACAAGGCAGAAGGCAGAAGGCAAAA
GTGAAGGCCAGAGAAAACTGGACACCACCTTTCCAGAGCACAGTTCAAAGGCAATGTCCT

CAAAGAAGACACTCCACCCTCCTCCCATTTCCTCCCTATTGCCTAAAAATAAGAAGGATA
[C,T]
GCGGCCTATGGCAAACCTTGGGCAGGCACGTGGGAGCTGAGCTCTTGCAAAGGGCAGATA
GTTCCTCTGGTGAGAGAGAAAAGGAAGGGCCAGTGAGGAGTGAAGGAAGAGACGAACAGA
GAGCCCGAAAGGCTGAGAACGTTGTCTGGCTTCCTGAAAGGCTTAAGGGGTTAGCTCTGG
AGGGTGAACTAAAAGCCCTAGTTATATTAAACACACACGCACACACGCACGCACACACAT
GCGCGCACACACACACACACATACACACAGTTGAAGGAGACCTGCAGTTTCCAAAAACAA

96000 TATTATTGATCTTGATTACTGAGTTTTTAGGTGTACCCTTAAATGTTGCACCTCTGACTT
ACTAGTCTCACCCTGATCCCTGTCCTGGATCTATGCCTGTCTGTTCTATATCAGCCTCTT
GCTTTGACCATAAGAATAACTTCAGACCTTTAAGCATAGAGGAAATAGGATTTCTGTCTC
CCTTCCCCACCTTTGTGATAATCTCAGCTTCTGCTTTTAAAGTCTATCTCCCAAGTAGTT
TGCCTACTATGTTCCTCCCAAGGTCACTAGGTTCTGTGAAACTAGCAGCAGGCTAGATTG
[T,C]
CACATTAGCACAAAGGATCCACTATTCCTGCAGCCGAGCTGGGACAAGCACTTAGGCCCA
CTGACTCCAACCCTTCAATAGCCTGGGACCTACGTTGTCTCCAGGTGGTATAAAACAAGA
ATTTCCCCTTTGACTGGGAGAAAAAGGGAAGAACTCTAAATTGGAAAACAGGTCATCTCG
AATTCTCACAGGTGGAAATTTCTGACAACCCCTTTGGGACCCACAATTCAACACACCCCA
AATGGGGACAGTAGCTAACATGCAACCTGTAGGCTGTTCTGTCATCCAGTGCCACTGTGC

96877 GGGAATTAAGGTGGAAGGCAGGGCGTTTTGACTGCATTTGACCCAAGTCCTGAAGAGCCA
GCTCCTCTCTCTTCCTAATTATTAGAAGGTTTTGTTTGGACCCAGTGTTTCACGTGTATA
CAATACAAACTTCTCTCTTTTCTACTTGGATCAAATTTGTTCTCTCAAAATAAGATTCCC
AGCAGTGAGAGAAGACAAGACAGAGAGATCCAACATCTCTAAAGCCATGAATCAGATAAC
CAGCCACTTGTTCTCTTCAGTGCTGGGAACAGATACACTGTTAAATAAAATGATTTTATA
[G,T]
ATTCTTCTCACTGCCTTTCCAAGAAGGGGATTTATCAACTTCAGGGCACAGCAATCATTT
ATTCCCAGACTACTGGCATGCATATATATATATATTTACTTCTCTTGACTTAGAAAAAAG
AGAGAATTGGAGTTGTGAATATTCCTGTCTCCCTCACCCCAGCCCCCTTGAAGTGAGTCA
GGACAAACTTGGGGCCCAAATGGAGCTGTAAGTAACTGAGTCACATGCAGAGATGAAACC
TTCACAGACCCACTGATATGGAGGTTGAAGATTAAATTCCCCTTTGAGAATAACTGGGTA

97271 ATTTACTTCTCTTGACTTAGAAAAAAGAGAGAATTGGAGTTGTGAATATTCCTGTCTCCC
TCACCCCAGCCCCCTTGAAGTGAGTCAGGACAAACTTGGGGCCCAAATGGAGCTGTAAGT
AACTGAGTCACATGCAGAGATGAAACCTTCACAGACCCACTGATATGGAGGTTGAAGATT
AAATTCCCCTTTGAGAATAACTGGGTAACACTCATACAGAGACTACTTTCAAGAAGGCCA
GATCCTCCCTCTAATGTATAGTGCAACGTTCCTAACCCTCAGCCCACTCCGTCATACCCC
[A,C]
ACTCACATGAATACACACATAAGCAGTAATATAAAGCACTTCCCACCATAGGGCAGCAAA
GAAGGAGGGAAATCTTTATTATGGAAGAGTGGAAGGAAGGAAGGAAGGGAAGGGAAGGG
AAGGGTAAGAGGAAGAATTCTCAGGGTGAGCAGAGGAATGACATGTTTGGGGCATAATGA
AGATAATTGAAGTGCAGAGTTTGTATGGAAAAAATTTGAAAATATCAGGTGGCAGGCCAGG
CATGGTAGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCCAAAGCAGGCGGATCACCTG

97470 ACTGGGTAACACTCATACAGAGACTACTTTCAAGAAGGCCAGATCCTCCCTCTAATGTAT
AGTGCAACGTTCCTAACCCTCAGCCCACTCCGTCATACCCCCACTCACATGAATACACAC
ATAAGCAGTAATATAAAGCACTTCCCACCATAGGGCAGCAAAGAAGGAGGGAAATCTTTA
TTATGGAAGAGTGGAAGGAAGGAAGGGAAGGGAAGGGAAGGGAAGGGTAAGAGGAAGAAT
TCTCAGGGTGAGCAGAGGAATGACATGTTTGGGGCATAATGAAGATAATTGAAGTGCAGA
[G,T]
TTTGTATGGAAAAAATTTGAAAATATCAGGTGGCAGGCCAGGCATGGTAGCTCATGCCTGT
AATCCCAGCACTTTGGGAGGCCAAAGCAGGCGGATCACCTGAGGTCACGAGTTTGAGACT
AGCCGGGCCAACATGGCAAAACCCCATCTCGACTAAAAATACAAAAATTAGCTGGGTTTA
GTGGCGCATGCCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCATTTGAGCC
TGGGAGGCAAAGGTTGCAGTGAGTCGAGATCATGCTACTACACTTCAGCCTGGGTGAGAG

97518 CCTCTAATGTATAGTGCAACGTTCCTAACCCTCAGCCCACTCCGTCATACCCCCACTCAC
ATGAATACACACATAAGCAGTAATATAAAGCACTTCCCACCATAGGGCAGCAAAGAAGGA
GGGAAATCTTTATTATGGAAGAGTGGAAGGAAGGAAGGGAAGGGAAGGGAAGGGAAGGGT
AAGAGGAAGAATTCTCAGGGTGAGCAGAGGAATGACATGTTTGGGGCATAATGAAGATAA
TTGAAGTGCAGAGTTTGTATGGAAAAAATTTGAAAATATCAGGTGGCAGGCCAGGCATGGT
[G,A]
GCTCATGCCTGTAATCCCAGCACTTTGGGAGGCCAAAGCAGGCGGATCACCTGAGGTCAC
GAGTTTGAGACTAGCCGGGCCAACATGGCAAAACCCCATCTCGACTAAAAATACAAAAAT
TAGCTGGGTTTAGTGGCGCATGCCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGGAGA
ATCATTTGAGCCTGGGAGGCAAAGGTTGCAGTGAGTCGAGATCATGCTACTACACTTCAG
CCTGGGTGAGAGAGCTTTCTTTTTTTTCTCTCACAAAAAAAGAAAAGTTCAGGTTGCAGA

98476 TGTCCCTTTCCCTCTAGCCACAGGTAACACGCTCTCCAGGCACTGGGAAAGTGGGTAATT
AGGAAAGCAGAGGAGTACCCATGGGCTGTGATGCCCAGTTATAAACCCAGACATTTCAGA
ATTAACAGAATGAGCATCAAGTCCTCAAATGGGTCTACATCCATAAACATGTCCAGCAGT
CAGCTCTTTACTGTCAGTAGAGACAAAATGTTCCTACACTTTCCCTAGGGGAAGCCACAT
CCTCAGTAGGTTATCTCTGATGAGTCCAGCTAGTCACAGGTATGTAGAAGCTGCATGCAG
[C,T]
AGAGGGCTCAAAGGAGGGTCCAGAATAGATACCAAAGCAAAAGGGGAGTCTGTGCACGTT
CTCACACGCACCCCGAAACACTCTTTTTGTTCACAAAATAGATGGTGTAGGGTAGTTCCA
AGAGATCATTTAGCTCAGGTTCCTGCCTCCATAAAATAAATAAGCCTTCCATATTAGTTG

TCTGTTGCTGTGTAGCAAATTGTCAGAAACGTAGAGGCTTAAAGCAATACCCATTTATTA
TCTCGCAAGTTCTGTATCTCAGAAGTCCAGGCAGGCTTGACTGGGTTCTCTGTCCAAGTT

98779 AGGGCTCAAAGGAGGGTCCAGAATAGATACCAAAGCAAAAGGGGAGTCTGTGCACGTTCT
CACACGCACCCCGAAACACTCTTTTTGTTCACAAAATAGATGGTGTAGGGTAGTTCCAAG
AGATCATTTAGCTCAGGTTCCTGCCTCCATAAAATAAATAAGCCTTCCATATTAGTTGTC
TGTTGCTGTGTAGCAAATTGTCAGAAACGTAGAGGCTTAAAGCAATACCCATTTATTATC
TCGCAAGTTCTGTATCTCAGAAGTCCAGGCAGGCTTGACTGGGTTCTCTGTCCAAGTTCT
[C,T]
GTGAGACTGAAATCAAGGTGTTGGCCAGGCTGGGATCTTATCTGGAGGCTCTGAGGACAT
ATACGCTTCCAACCTTATTCAGGCCATCAGCAGAATCCCGTCTCTTGTGGCTTGAGGTTG
GAGGTCCCCGTTTCCTTGCTGGCTGTCATCCAGGGACCACTCTTTGCACCTACAGGCTGC
CTATGTTCCTATTCACAAGACACCGTTCATCTTCAAACCAAAGCAGCATGTAGAATCTTT
CTTGTGGCTCGTATCTTTCTGGCTTTCCCTTCTTCTTTAGCCAGAGAAAGTTCTTTGCTT

99218 CTGGCTGTCATCCAGGGACCACTCTTTGCACCTACAGGCTGCCTATGTTCCTATTCACAA
GACACCGTTCATCTTCAAACCAAAGCAGCATGTAGAATCTTTCTTGTGGCTCGTATCTTT
CTGGCTTTCCCTTCTTCTTTAGCCAGAGAAAGTTCTTTGCTTTTAAGCGTTCATGCGATT
CAATCAGGCCCACCTGGATAATGTCCCTATTTTAAAGGTAACTGTGATACCGTATAACAT
TTCAGGAGTGATAACAGCACATTTACAGGTTCCAAGGATTGGGGCAGAACATCTTTGGGG
[C,G]
AACATTTTAGAAACTCTGCCTCCCCACTCACCCATAATCCTTTTAAAAACCAAATCTTGA
AGCCTTTTTTTCCCAAAGGCCTTTTTGAATAAGCACATTTATACCTAACTTCATCAGACA
CCCACTTTGAGCAAACACTAGCATGTGGCAAAATAGGCTGTAAATCAATCAGAACTATTC
TTTCCCACCACAATCTTTCTCAAACACATTGGGAGAATCTGACACTGTCAGTGGTATACC
AGAGCAGACTCCTACCATCTCACAAGAGCTGACTGTTAAATGTTTAGTAATTGTGGACAT

100538 TGTAACTATTGGTAAGTTAATTTGAAATGTGGTTTCTAGATCTCTCATCATCCTAGTCAC
CCTACTCTGGATGTACTCCAAAGTCCCTCTCAAGATATAGTGTCAGAATTGACCTAATTA
GTCCAGCATTTGACTGAAACGCTAGACTTTGACTCCAGCCCCCCATCCTTGACTGGCACT
AGCATTCAAGCCGCTTCTCCTCTTTCCCTGGGTCTTTAATAGAGTCAGAGCGACTTCTCC
AGGGGATCTTTTGGCCATGGACCAGTAGCATCCACACACGCTGGGGCCTTGTTAAAAAGG
[C,A]
AGGCTCTCAGGCCCCACCCCAGATCTACTGAATCAGAATCCACACATTAACAAGATGCTT
GGGTGATTCATGTGCACATTAAAGTTTGAGAAGCACCGCTTTCAGGGACGAGATGACACA
CTTATTTTAAAGAGAACGCCAATTAGAGACCCTAAGCCTTCTCATGGAACAGGGGCCTTC
CCCTCAGACCTTGGGAGAGGGGTCAGGGAAATATCAGTGTTGGGTTGTTGGTGACAGGTG
GCGGTGGGGGGTTCAGTCCACGTTCAAAGAGCCAGAAACCTGGCAGGGGAAGAGATGGGG

101045 GGAAATATCAGTGTTGGGTTGTTGGTGACAGGTGGCGGTGGGGGGTTCAGTCCACGTTCA
AAGAGCCAGAAACCTGGCAGGGGAAGAGATGGGGCAGTGACACCCAACCGGAAAAATAAA
GGAAACTACAAGAAGAACCCAGCTAAGAGATGTGAGGCTTCTGAAAGCTCCCATGGAAAG
GTTCGCAGCTCCTCCACCTGCTCGGTCCAGCTGCCCCAGGTCAAGGAAGCTCTGTGAGTG
TTAGCTGACCCGGAGCAGCAAGGATACATTCAGAAGTGATGAAAGGGAACGCTTCTTGAC
[A,C]
GGGTAAAGAGTCATTCAGTAGGAATGAGACAGGAAGAGGTCACAGAGTCAGAAGCCCAGC
CTGTACTCAGAGATTATTTCTGGCATGGGAGGGCCGAAGGGTTAGGAGGCCACCTACTCA
CAATACAATACAGAGGCAGATCCACTTATTACCTGCCTGTGCTGCTGGGATTTCAGTGTG
GAAATTCTGTGCCTCCTCACTGTGGCTGCAGCTTGGGAATGACATCCAGAGCTTACCCAC
CTGCATAAGAAATAAGCTATAGGTGTAATAGGGGGACATAGGCTAAAATCCTAGCTCAGC

101232 GCTCCTCCACCTGCTCGGTCCAGCTGCCCCAGGTCAAGGAAGCTCTGTGAGTGTTAGCTG
ACCCGGAGCAGCAAGGATACATTCAGAAGTGATGAAAGGGAACGCTTCTTGACAGGGTAA
AGAGTCATTCAGTAGGAATGAGACAGGAAGAGGTCACAGAGTCAGAAGCCCAGCCTGTAC
TCAGAGATTATTTCTGGCATGGGAGGGCCGAAGGGTTAGGAGGCCACCTACTCACAATAC
AATACAGAGGCAGATCCACTTATTACCTGCCTGTGCTGCTGGGATTTCAGTGTGGAAATT
[C,G]
TGTGCCTCCTCACTGTGGCTGCAGCTTGGGAATGACATCCAGAGCTTACCCACCTGCATA
AGAAATAAGCTATAGGTGTAATAGGGGGACATAGGCTAAAATCCTAGCTCAGCTGCTTAA
TAGCTGTGCGACTGAGCAAGTTACTTAACCTCTTTGAGCATCTGTTTTCTCATCTTTAAA
ATGGAAGTAATCATAATTGACCAGGCCCAGTGGCTCACACCTATAATCCCAGCACCTTGG
AAGGCCGAGGCCAGTGGATTGCTTGAGCCCAAGAGTTTGAGACCAGCATGGTGACACCTC

101266 CAAGGAAGCTCTGTGAGTGTTAGCTGACCCGGAGCAGCAAGGATACATTCAGAAGTGATG
AAAGGGAACGCTTCTTGACAGGGTAAAGAGTCATTCAGTAGGAATGAGACAGGAAGAGGT
CACAGAGTCAGAAGCCCAGCCTGTACTCAGAGATTATTTCTGGCATGGGAGGGCCGAAGG
GTTAGGAGGCCACCTACTCACAATACAATACAGAGGCAGATCCACTTATTACCTGCCTGT
GCTGCTGGGATTTCAGTGTGGAAATTCTGTGCCTCCTCACTGTGGCTGCAGCTTGGGAAT
[G,A]
ACATCCAGAGCTTACCCACCTGCATAAGAAATAAGCTATAGGTGTAATAGGGGGACATAG
GCTAAAATCCTAGCTCAGCTGCTTAATAGCTGTGCGACTGAGCAAGTTACTTAACCTCTT
TGAGCATCTGTTTTCTCATCTTTAAAATGGAAGTAATCATAATTGACCAGGCCCAGTGGC
TCACACCTATAATCCCAGCACCTTGGAAGGCCGAGGCCAGTGGATTGCTTGAGCCCAAGA
GTTTGAGACCAGCATGGTGACACCTCGTCTCTAGAAAAAATACAAAAATTAGCCAGGCAT

101290 TGACCCGGAGCAGCAAGGATACATTCAGAAGTGATGAAAGGGAACGCTTCTTGACAGGGT
AAAGAGTCATTCAGTAGGAATGAGACAGGAAGAGGTCACAGAGTCAGAAGCCCAGCCTGT

ACTCAGAGATTATTTCTGGCATGGGAGGGCCGAAGGGTTAGGAGGCCACCTACTCACAAT
ACAATACAGAGGCAGATCCACTTATTACCTGCCTGTGCTGCTGGGATTTCAGTGTGGAAA
TTCTGTGCCTCCTCACTGTGGCTGCAGCTTGGGAATGACATCCAGAGCTTACCCACCTGC
[A,G]
TAAGAAATAAGCTATAGGTGTAATAGGGGGACATAGGCTAAAATCCTAGCTCAGCTGCTT
AATAGCTGTGCGACTGAGCAAGTTACTTAACCTCTTTGAGCATCTGTTTTCTCATCTTTA
AAATGGAAGTAATCATAATTGACCAGGCCCAGTGGCTCACACCTATAATCCCAGCACCTT
GGAAGGCCGAGGCCAGTGGATTGCTTGAGCCCAAGAGTTTGAGACCAGCATGGTGACACC
TCGTCTCTAGAAAAAATACAAAAATTAGCCAGGCATGGTGGCAGGTGCCTGTAGTCTTAG

101326 AAAGGGAACGCTTCTTGACAGGGTAAAGAGTCATTCAGTAGGAATGAGACAGGAAGAGGT
CACAGAGTCAGAAGCCCAGCCTGTACTCAGAGATTATTTCTGGCATGGGAGGGCCGAAGG
GTTAGGAGGCCACCTACTCACAATACAATACAGAGGCAGATCCACTTATTACCTGCCTGT
GCTGCTGGGATTTCAGTGTGGAAATTCTGTGCCTCCTCACTGTGGCTGCAGCTTGGGAAT
GACATCCAGAGCTTACCCACCTGCATAAGAAATAAGCTATAGGTGTAATAGGGGGACATA
[G,A]
GCTAAAATCCTAGCTCAGCTGCTTAATAGCTGTGCGACTGAGCAAGTTACTTAACCTCTT
TGAGCATCTGTTTTCTCATCTTTAAAATGGAAGTAATCATAATTGACCAGGCCCAGTGGC
TCACACCTATAATCCCAGCACCTTGGAAGGCCGAGGCCAGTGGATTGCTTGAGCCCAAGA
GTTTGAGACCAGCATGGTGACACCTCGTCTCTAGAAAAAATACAAAAATTAGCCAGGCAT
GGTGGCAGGTGCCTGTAGTCTTAGCTACTCGGTAGGCTGAGGTGGGAAGATTATATGAGC

102342 ACCCTGTCTCAATAAATAAATAAGAAGAATGAAACAAGAAAGTTCTTCTTATGGTTCTCA
TGGTGGTGAGCACAATGTAAGCATATATATTATCTTAGAATTCTTCCTTCCTGTATAAAG
AAGGCCTCCTCCAATGTATTAATCATCTGTTCAACTAATAAATGCTGCTTACTCCCACTT
TCACTCTAAAGGAACTCAATGGCTAAAGAGAACCCCTTCCCCTTTGCAGCACCCTGAGGAT
CAGAGGCCTGATTTGAATGTCCTCGATGCAAAGGACTATTTCAAAAGGCCAGCCAGGCAG
[C,A]
CCAGACATGTATTTCCTAATCGTCTCCAGGTTGTTTGATAGAAGATCTCCTGGGAGCAGG
TTTCCGCAGCAGCTCAGCCAGGTCTGTTCTGGGAACGCTGTGTGCATTGGCACCTCCCTT
GGCAGAAAGCTTGGAGGAAAGGCAGGTGCAGGTCCTGGAGCCTCTGACAGCATTACTGGC
TCTAGGAGTAGCTGCTCAGGATAATCTGTCCCCATGACCATTAAGTAACTGCCACTGTGC
GGGAAGAAGAACTGGAAATGGGGGGCCCAAAAAAATCTGAAAACCCTCACTTGAACCAGT

104489 GTTCAAGAGCTGGAAGGGATTTTTCTAGCCTCCAGGCAAGGTAATACCATAAGTCCCAAC
AGTGATGCCCTCCCTGGGAATGATCTCAATGGGAGAATCCTATACCCTGCCTCCTCCATT
CATTCCTTGCTCTGATGGTGGTTCTGGCTGGCTAACCTAAGTTACTCTTGCCACTAGTTA
ACGCCTGTCCTTATTTCTCTTGTCCCCACCTAAGATGTCAATCAAAACAGCACGAGCCAT
GCTATGTCACATGACATGTTGTCTGTCCAGCCCAGAGCTTGTTGCTGATGGGGGCACAGA
[C,T]
TAGATTTTGAGAGAAATCTCTCTGTTACCACCCTTAACATTCCAACCCCCTCTAATAGCC
CATTTAGGATTTATCATACTGTTTCATCCAAACCTTTCATGACCTGATTTCTATTTCCAG
CTTCAACCACCCCTTGGGTCACCACCTGTACTTATTGAGTTTCCCTAGTTTTCTGAATTA
ATGACTGAAGATGATAAGCTTCCCTTACATATGACTCTCAAACCACCAAACTGGGATTGT
TGTTACTCTTAGTGATAATGGTTGCTATTTATGAAACTTTTAATAGGGAACACAAACCCT

105266 AGGCCAGAGCATCATGGCCTTTCACAAGTTGAAGAGCCACGGGCTTTCTACGGTAGCCAG
CCACGCTTTTCCATGACTGGGGTGGGTGTGGCAAGTGATGAGGGTTTGGAGTTCATGTGG
TGGGGTGGCAGGGACCAGGTGTCTTGGTAACTGCTGTTGCATTCACTTCAGGAGCAAAGG
ACCAGATCTGATTCTGCAGGATCAACAATATGGACACTGCAGGCTCTGTAGACATCCAAA
GCTCTAATGGTGACTTGGGGAAGCTCAGGAGGGCAGGGAGGTTGTACCCATTTAGAATGT
[A,G]
AAGATTCCTATTTTATAAAAAAGAAAAAAAGGAGACTGAAGGCCTCAGTCTCCTCCAACA
AAGCCAGGCTGTGGGGTAGCAGAGTCTCAAAGGGTGCAGGCCCATGGCCACTGCCCAGGG
CTCCTGCTCAGGCCTCCTCACTCCCACAACTGAGGGGAGACCCAGTTCCACACCCACCCA
CCTAGCAGTGTCTCACACCCACCGGGAGAGGTCTAAACATCTTCCCTGGGAAATGGTCCC
AAAATGTCCCTGCAGTAAGCAACCATCTGGAGAGGCCCAGGTCTACATCTGTTTTTAAAG

105338 ATGACTGGGGTGGGTGTGGCAAGTGATGAGGGTTTGGAGTTCATGTGGTGGGGTGGCAGG
GACCAGGTGTCTTGGTAACTGCTGTTGCATTCACTTCAGGAGCAAAGGACCAGATCTGAT
TCTGCAGGATCAACAATATGGACACTGCAGGCTCTGTAGACATCCAAAGCTCTAATGGTG
ACTTGGGGAAGCTCAGGAGGGCAGGGAGGTTGTACCCATTTAGAATGTAAAGATTCCTAT
TTTATAAAAAAGAAAAAAAGGAGACTGAAGGCCTCAGTCTCCTCCAACAAAGCCAGGCTG
[T,C]
GGGGTAGCAGAGTCTCAAAGGGTGCAGGCCCATGGCCACTGCCCAGGGCTCCTGCTCAGG
CCTCCTCACTCCCACAACTGAGGGGAGACCCAGTTCCACACCCACCCACCTAGCAGTGTC
TCACACCCACCGGGAGAGGTCTAAACATCTTCCCTGGGAAATGGTCCCAAAATGTCCCTG
CAGTAAGCAACCATCTGGAGAGGCCCAGGTCTACATCTGTTTTTAAAGCTCCAATAAATA
AATAAATGAAGGAAGAAAAAAAGAAGAAGAAATGCAGAACAGGGTGACTAAAATTGGCAT

105570 ATTCCTATTTTATAAAAAAGAAAAAAAGGAGACTGAAGGCCTCAGTCTCCTCCAACAAAG
CCAGGCTGTGGGGTAGCAGAGTCTCAAAGGGTGCAGGCCCATGGCCACTGCCCAGGGCTC
CTGCTCAGGCCTCCTCACTCCCACAACTGAGGGGAGACCCAGTTCCACACCCACCCACCT
AGCAGTGTCTCACACCCACCGGGAGAGGTCTAAACATCTTCCCTGGGAAATGGTCCCAAA
ATGTCCCTGCAGTAAGCAACCATCTGGAGAGGCCCAGGTCTACATCTGTTTTTAAAGCTC
[C,A]
AATAAATAAATAAATGAAGGAAGAAAAAAAGAAGAAGAAATGCAGAACAGGGTGACTAAA

ATTGGCATGTATTTTTAAATGTTTATATTAACAAACTAACACCTTTTAACATGAAAAGCA
ATATAATTGTGCTAGCCACAAAATCATCGTAGGACTGAGAAAGGAATCGTGATTCTGAGA
GCCCTAGAGTTAATGTGATCCAGCTGGCTCATCCCTGTGACTGCAGAAGCCTGTTTGGAG
ATAGTGTCAGTAGCTTTTCAGGCCCTCTGTGAATTGCCAGAATGTGTGACATGAGCCAAA

105928 AAAATTGGCATGTATTTTTAAATGTTTATATTAACAAACTAACACCTTTTAACATGAAAA
GCAATATAATTGTGCTAGCCACAAAATCATCGTAGGACTGAGAAAGGAATCGTGATTCTG
AGAGCCCTAGAGTTAATGTGATCCAGCTGGCTCATCCCTGTGACTGCAGAAGCCTGTTTG
GAGATAGTGTCAGTAGCTTTTCAGGCCCTCTGTGAATTGCCAGAATGTGTGACATGAGCC
AAATTTCCCCCCAGCATCCCCGCCGCCGCCACCACCACCCCCGACCCAACCCTCCCGCCG
[G,A]
CTCCCATAGAATAGTCACTGCCATACAGAAAAAGAGAAGTTCTACTATTTCTGGGCAAGA
TTTCCACAAACCAGTTTGTCCCTTTCTGCTTTCATGAAATAAACCATTTGGATCAACGTC
AGCTGATTGCAAAAATTTTCCCTTGTCTCAAAAGCAAGACTGATAAGGAAGCAAACATGG
GAGGACCTTAGTGGCCGAGCCTTTATGTGTATGTTATTTCATTGCTCTCATAACTGCCCT
GGGATGCTGTAAGCATGATTCATCCTGTTTGTTTATCAGTTAAATTATGTATCCAAGATT

106459 TAACTGCCCTGGGATGCTGTAAGCATGATTCATCCTGTTTGTTTATCAGTTAAATTATGT
ATCCAAGATTACACAGCCTATCCAGGATTAGAACTCAGAGCCCTCGGCTGTGAAGCTTGA
GCTCTTTCTTTTCAGTCTTCAAATATGATCATGCCATGAAGCAGCACAAAGCCCAGGAGG
AGCCCAGTGAGGCTGGAGGGGTCCACTGGCAGCCACTCTCCTCCGTGCCCCTGTGGTGTT
GGGGCAAACTTGGATCTTTCTGAATCTTTTAACTGTTTCCTTCTCTTCCCGTTTTTGTCT
[G,C]
CTGGCTGACTTGTCCTACACTCTACTCCTTGCTTATGATACTTATTTTTCCATCCACAGC
AAAACAATTCACATCAAGGTAATTGATGATGAGGCATATGAGAAAAACAAGAATTACTTC
ATTGAGATGATGGGCCCCCGCATGGTGGATATGAGTTTTCAGAAAGGTGTAGTACCCTGT
CCTCCACACTAACACTAACATTCTTCTCTCCTCTTCTGTTTCTTCCTCTCCAACCCATTT
GTCTCCTCCTCCTCTTGTCTTCCACCTCTCTGGTTCCCTTTCCCTTGTCTCCTCTCTTGC

107710 CTTCAATGACCCCATACATCCCATGGCCTCCAATAGACAAGTCAAGAAGTCCTTTCCTGA
ATAGATCATACTGTGGAGCAGGGAGCTGCCAGTACTGAGGGCAATGTTCCTTCCCCTTCC
AAGCTGTCCCTCATGCCCTCCAGTACATGCCTGTTGTCACAGAGCACCCCAATCCCATCC
CACAGCAGAGTTCCTGCAGCAGAGAAACAGGCTCACACCTTGTAGACAGCCCTGGGGTCC
CATATCTAGGGCCAACAGAAATATTCCCAAAAAAATGCCTCTTGACAATCAATGAGCTTT
[C,G]
TCTTTTGTCCGCTGAGCAAGGTATAAAAAGATGTCAAAAGAAGTACCCAAAAAGGTAATA
AAAATGTACAGTCGTGCATCACTTAGCAATAAGGATACATTCTGAGGAAGGTGTCCTTAA
GCAATTTTGTCATCGTGGGAAAATTATAGAGTGTACTTTCACAAACCTAGATGGTGTAGC
CTACAACACACCTGGACTATGTGGGCCTATTGCTCCTAGGCTACAAACCTGTACAGCATG
TGCTTGTACTGAATATTGCAGGCAACTGTAGCACAATGGTATTTGTGTATCTAAACACAT

108062 AAGGTAATAAAAATGTACAGTCGTGCATCACTTAGCAATAAGGATACATTCTGAGGAAGG
TGTCCTTAAGCAATTTTGTCATCGTGGGAAAATTATAGAGTGTACTTTCACAAACCTAGA
TGGTGTAGCCTACAACACACCTGGACTATGTGGGCCTATTGCTCCTAGGCTACAAACCTG
TACAGCATGTGCTTGTACTGAATATTGCAGGCAACTGTAGCACAATGGTATTTGTGTATC
TAAACACATCTAGACATAGAAAAGGCACAGTAAAAATATCGTAGTATATAGCCTTATGGG
[G,A]
CCACTATTGTAGATGTGGTCTGTCATTGAGCAAAACGTTTTTATGTAGCATGTGACTGTA
CTTGTAAAGTACACACACCACAAATGCACAGCAAGTCCTGTGCCCTACAAGCCCCTTTGG
GTCAGTCTACTACATTATAAATGGCAAAGCCGAGCACGCCCACAGAAGGTAGCAGGAACA
TCAGAGGATCTGAAGAGACATTTAGGTAAATGCTCTTTACCCTTTAGAGCATTTAGTTCT
TAGGCCTCCCCTCCCCCAATCTCCCCCCCGCCCCCCGCCAAAAAGAAAAAGAAAAAGAAA

108214 GGCCTATTGCTCCTAGGCTACAAACCTGTACAGCATGTGCTTGTACTGAATATTGCAGGC
AACTGTAGCACAATGGTATTTGTGTATCTAAACACATCTAGACATAGAAAAGGCACAGTA
AAAATATCGTAGTATATAGCCTTATGGGACCACTATTGTAGATGTGGTCTGTCATTGAGC
AAAACGTTTTTATGTAGCATGTGACTGTACTTGTAAAGTACACACACCACAAATGCACAG
CAAGTCCTGTGCCCTACAAGCCCCTTTGGGTCAGTCTACTACATTATAAATGGCAAAGCC
[G,A]
AGCACGCCCACAGAAGGTAGCAGGAACATCAGAGGATCTGAAGAGACATTTAGGTAAATG
CTCTTTACCCTTTAGAGCATTTAGTTCTTAGGCCTCCCCTCCCCCAATCTCCCCCCCGCC
CCCCGCCAAAAAGAAAAAGAAAAAGAAAGCAGAAAATTACAATTCTGGCTCACTAGTAGG
ACCTGCTAGCCACCATTGTGATTCCATGAAGGACCAGAAGAAACCATATAGGAAGAATCA
GGCCCACACGGCAACCTCTCCACATGACAAAGAGCCAGTCTTTGGAGGGCAGTGAATTTC

108364 CACTATTGTAGATGTGGTCTGTCATTGAGCAAAACGTTTTTATGTAGCATGTGACTGTAC
TTGTAAAGTACACACACCACAAATGCACAGCAAGTCCTGTGCCCTACAAGCCCCTTTGGG
TCAGTCTACTACATTATAAATGGCAAAGCCGAGCACGCCCACAGAAGGTAGCAGGAACAT
CAGAGGATCTGAAGAGACATTTAGGTAAATGCTCTTTACCCTTTAGAGCATTTAGTTCTT
AGGCCTCCCCTCCCCCAATCTCCCCCCCGCCCCCCGCCAAAAAGAAAAAGAAAAAGAAAG
[C,A]
AGAAAATTACAATTCTGGCTCACTAGTAGGACCTGCTAGCCACCATTGTGATTCCATGAA
GGACCAGAAGAAACCATATAGGAAGAATCAGGCCCACACGGCAACCTCTCCACATGACAA
AGAGCCAGTCTTTGGAGGGCAGTGAATTTCAAGGAAAGTTTTCTTCCCTGGGTGACTTGT
TTTTAAAAGATGTTATGTTTTGTTGAGATACCCAGAGATGAACAGAAACTTCCATCACCT
TGTGCCCCAGACCCATGATAATTCACATTGAGGAAACCAGTTTTGGAACACATCACCCCT

108657 AAGAAAGCAGAAAATTACAATTCTGGCTCACTAGTAGGACCTGCTAGCCACCATTGTGAT
TCCATGAAGGACCAGAAGAAACCATATAGGAAGAATCAGGCCCACACGGCAACCTCTCCA
CATGACAAAGAGCCAGTCTTTGGAGGGCAGTGAATTTCAAGGAAACTTTTCTTCCCTGGG
TGACTTGTTTTTAAAAGATGTTATGTTTTGTTGAGATACCCAGAGATGAACAGAAACTTC
CATCACCTTGTGCCCCAGACCCATGATAATTCACATTGAGGAAACCAGTTTTGGAACACA
[T,A]
CACCCCTAAGTGATAGAAGCCCAAAGGTGATTTAGAATTTGATGATTTACATCATTTTCT
TCACATTTTCCCAGAAATGCATCAGCTGTAAATAGTAAAGGATTCCTATGTAATATTGTG
GTTAATACATATTTATTTTAGTTCCCACCACTGAAGCCCTATGAGATAAAGAATGAGAAA
GATCACACAATTCTACCTCCCTTTCTTCTCTCTCTCTCTCTCTTTCTCTTTCTCTCTCAC
TCTCTCTCTCTCTCTCTTCTCTTCCTCTGTCTGGTTTTCCTTCCTCATAAATACTTTT

109746 GCTCCAGGTGACCAACCAATCTACCGACCCAGTCGACACACTCTCTCTCTTGTTGTCCCT
ACAGGAAAACCATAAGGGTTAAAATAGTAGATGAGGAGGAATACGAAAGGCAAGAGAATT
TCTTCATTGCCCTTGGTGAACCGAAATGGATGGAACGTGGAATATCAGGTGTGAGATTCT
TTAAAAACAAAACAACAAAAAAAAAAGAAAGAAAAATTAAAACAAACTGAAAAACAACAAC
AAAAAAGAAAAAGCAGCTATATTTTTGTCTCCCTCCTTTTCTTCCCTTCTCCTCCTTTCT
[C,T]
TTTTTGACCAATGGATTTTTTTATTCTTTTCCCTCCTGTATTCTCGCTCTCACCCTGTTT
CGGTATCATCTCTGCCTTCTTAGCCTTAGCTTATTCCAAATTCCTCCTTTACCGCCTTCT
GGGCAGCACTGCAGCCTCAACTCCTCATTACCCTAATGAGTTATTTCCCTGTTTTGCTAC
AATTTTCAATTATTCAATTGCCATGGGCCCCTGCACTCTCCCCCACCCCACCCCTACACT
GTAACCTGTAAATGTGAAAATTCCTTGGTGGGTGGGGAGGAGAAGAAAAAAAAAGGAATGT

111484 ACAGGCCTTCTCAGTGTGATTGGTCATTTCTCATTGTCTGCTGGGGACTCTCCTGCAGAG
CTGACCACTTCTGTGCCTGCGCTGGTTTGGACACACCTGATGCTCTAGGGGCAGAACTCC
TCTCCTTCTTCACTGCTGGTTCTCTTCGTCACCACTCAATAAAACGTTGCCCTCAGCCTG
ACTGCCAAAAAGTGCTGGAAGAAAGAAATTATCTCTGGTTCTATTGTTTCCCACATTGTA
TTCTTGCCCAACTTCCAGTTCTTGCCACCAACAATATTCTCAGAGGTTGCCTCAGCACCT
[G,T]
CCCTACCTCATTCCCACCTCCCTTGAGCATTTATTCCATGTATTCATAATTGGTTGGAAG
CAGCAGATACCCAAGGCCAATTGTAAGTCACCTTCATCAGTTTCCACAGTCCAAGCTACT
TAGATGCAAACGAAAGCAGCACATGTACAGCGTACAGGAAGGAAGGCAGTGGTTCCAGAC
AAGAGGAAGAGATTGGAAGTCCATACATGCCTTTATTCCACCAGTAAAAAGGCTCTTCTC
TTATGCCTCCCTTAAAACCTCTACCAACAGCAGGACAGAGAGTGACCCAAGATAAGTCTT

112879 TTGGTCCTCCTTTTTCGACAACCGTGGGCTCATCTTGACAAGCTGCCCAGATGCTTCCTA
ATTACTCACAGTCCTATGCTCTTTCCAGCTTGTCCCTGGGGTGTCTGAGCAGGAATAAAT
GACTCTCACCTGACCCAGGGGATCAATACAGGGGAAAGTTCAGCTCCAGCTTCTCTCATG
AGCAGCAGCAGGAAAAACACCCTCGAGGTATTGTGTCAGTCAAAGCTGGCCTACCCAGGT
CTTGCTGACCCATCTATAACTGCTGAGCAGAAAGTCTTGGATTCATGGAGACAATGACCA
[A,G]
AGAATGATGGAATTCCAGCCAACTGCAGGCCTTCTCACTACTCTAGGGATGGGCCAGATG
TTCGGTGGCATGTATGAGTGAAAACCAGGGCATCAGGGACCTTTCTGGAAGAGCTGCCTT
TGTCTGACCCACCTGTGTTCATTTATGTGCTGGGATCTCTGATCTCCCCTGGAACTTGGG
GGAAGCTCTTCCACGCAAACTCCCGGAAGGAGCAGAATAAACAAGCTCTTGCCTATCTAT
CTATCTATCTATCTATCTATCTATCTATCTATCTATCTACCTATCTGCCTATCTATATCT

113245 TGGCATGTATGAGTGAAAACCAGGGCATCAGGGACCTTTCTGGAAGAGCTGCCTTTGTCT
GACCCACCTGTGTTCATTTATGTGCTGGGATCTCTGATCTCCCCTGGAACTTGGGGGAAG
CTCTTCCACGCAAACTCCCGGAAGGAGCAGAATAAACAAGCTCTTGCCTATCTATCTATC
TATCTATCTATCTATCTATCTATCTATCTATCTACCTATCTGCCTATCTATATCTATCTA
TCTCAATGTAGTGAGGAAAGCCATTGATCCATTAACCTTTGGAATTCTACATGGGAGATA
[C,T]
CTAAAAAAGTGAACTGCCTTGTTTATGTATCATGCAGACTCTGGATCCACATATATCTCA
GTGGCTGTGAATATAGGATGATTGATCACAGGCCTGAGTTGCATTCCTACAGATTCTTAG
GAAAAAAATTGATTCACAGACATGTCCCCCCTGGTTCCCCCACAACACACACTCCTTCCT
CAGCAATCTCTATCAGTCACCAACTACACGTTGAATATGTGGCAAGCTCTTCCCAGACCT
TTATCTGAGAGCCAAGGAGTGAGGGGCTGTACTAAGATATCATAGAAATGAAAATGTGGT

113265 CAGGGCATCAGGGACCTTTCTGGAAGAGCTGCCTTTGTCTGACCCACCTGTGTTCATTTA
TGTGCTGGGATCTCTGATCTCCCCTGGAACTTGGGGGAAGCTCTTCCACGCAAACTCCCG
GAAGGAGCAGAATAAACAAGCTCTTGCCTATCTATCTATCTATCTATCTATCTATCTATC
TATCTATCTATCTACCTATCTGCCTATCTATATCTATCTATCTCAATGTAGTGAGGAAAG
CCATTGATCCATTAACCTTTGGAATTCTACATGGGAGATACCTAAAAAAGTGAACTGCCT
[T,C]
GTTTATGTATCATGCAGACTCTGGATCCACATATATCTCAGTGGCTGTGAATATAGGATG
ATTGATCACAGGCCTGAGTTGCATTCCTACAGATTCTTAGGAAAAAAATTGATTCACAGA
CATGTCCCCCCTGGTTCCCCCACAACACACACTCCTTCCTCAGCAATCTCTATCAGTCAC
CAACTACACGTTGAATATGTGGCAAGCTCTTCCCAGACCTTTATCTGAGAGCCAAGGAGT
GAGGGGCTGTACTAAGATATCATAGAAATGAAAATGTGGTGTGTCACAAGTTTCCTTAAT

113497 GAGGAAAGCCATTGATCCATTAACCTTTGGAATTCTACATGGGAGATACCTAAAAAAGTG
AACTGCCTTGTTTATGTATCATGCAGACTCTGGATCCACATATATCTCAGTGGCTGTGAA
TATAGGATGATTGATCACAGGCCTGAGTTGCATTCCTACAGATTCTTAGGAAAAAAATTG
ATTCACAGACATGTCCCCCCTGGTTCCCCCACAACACACACTCCTTCCTCAGCAATCTCT
ATCAGTCACCAACTACACGTTGAATATGTGGCAAGCTCTTCCCAGACCTTTATCTGAGAG

[C,G]
CAAGGAGTGAGGGGCTGTACTAAGATATCATAGAAATGAAAATGTGGTGTGTCACAAGTT
TCCTTAATTCTTAGATCTTAAACTCTAAGAGGGTTCAGCATAAGTACAAATTCAAGGGCT
AGAGACAACCTGTATTGGGTGTGTCTTTAACTCAGTTTCCCAATCCACATAGGGACCTTG
CATTTGTCATCTCTCATCTATGTATAGCTGTTGGTATGACAGTTTCTCTGTTCCAGAATA
CCTGAACTCTGACTTAGCCTGTCCTTTCTGAAACAGAAAAATCACCCAACCAGAGATCTA

114486 CCCCATGGTCATTTTTGCCACTCATAAGTTAGCTACTCTGGCAGGGTTGCAACTTACACA
GTTTTCATGATAACTGGATTCTCACTCCTTTTTTTACAGAATGGATGTGATAACCTGGTA
TCCTACACAGTCATGAGTGACCAACCTACCCATTTGGTTCCCCATCCTCATTCCTCCATT
CCTAGCCCTAGGGTAGCCGGGAAAGCATAGGAGCAAATGCCCTTACCAGGGCCCTGGTGC
TCAGCAGCCTCTCCGGCTGCTCACACCTCTTGCTGCTGCTCTGTGCATGCTCCAAAGGCT
[G,T]
CTTTTTGCGTATGGCTGCTGAGCTCTCACCTACTAAGCTCTCTGCTTTCCTTATGCTGCC
AGCAACCACAAAACCTGGTGATACTTTCAAGATGGGACATTAATGCTCTTTCCTTTTCTT
TCTTCCATTTTTCTGGTATCCATTTGCAAACAGCGCTCCTGTTATCTCCAGGTAAGAGGT
GTCTTGTCCCCCTCTTTTCTTTCCACTTCTTGCCAGTGCCATTATTTGGTTTAAGACCAA
TGTCCTTTGATTTATTGAATAAGAACTGCAGGCTCAAGTTAACCTGACAATTTCTCCCAA

114686 GAAAGCATAGGAGCAAATGCCCTTACCAGGGCCCTGGTGCTCAGCAGCCTCTCCGGCTGC
TCACACCTCTTGCTGCTGCTCTGTGCATGCTCCAAAGGCTGCTTTTTGCGTATGGCTGCT
GAGCTCTCACCTACTAAGCTCTCTGCTTTCCTTATGCTGCCAGCAACCACAAAACCTGGT
GATACTTTCAAGATGGGACATTAATGCTCTTTCCTTTTCTTTCTTCCATTTTTCTGGTAT
CCATTTGCAAACAGCGCTCCTGTTATCTCCAGGTAAGAGGTGTCTTGTCCCCCTCTTTTC
[T,C]
TTCCACTTCTTGCCAGTGCCATTATTTGGTTTAAGACCAATGTCCTTTGATTTATTGAAT
AAGAACTGCAGGCTCAAGTTAACCTGACAATTTCTCCCAAGGACTGGGAGATTTATTTTC
CCACATGAAGCAATTATGAGAAAGCAATTGTGAGGAAGGCAATTCCTTGAGCATCACTTC
TGTCTGGGGACGTGGGTTAAGGCATAGCTGATCCTCTCTGGGACCAGGAAGAGAAATTAA
GCTTAACAAGGAGATGGTGGGTCATAGACTTCTCCTGAGTCTTAATTCATCTGCCATCTC

114817 TACTAAGCTCTCTGCTTTCCTTATGCTGCCAGCAACCACAAAACCTGGTGATACTTTCAA
GATGGGACATTAATGCTCTTTCCTTTTCTTTCTTCCATTTTTCTGGTATCCATTTGCAAA
CAGCGCTCCTGTTATCTCCAGGTAAGAGGTGTCTTGTCCCCCTCTTTTCTTTCCACTTCT
TGCCAGTGCCATTATTTGGTTTAAGACCAATGTCCTTTGATTTATTGAATAAGAACTGCA
GGCTCAAGTTAACCTGACAATTTCTCCCAAGGACTGGGAGATTTATTTTCCCACATGAAG
[C,A]
AATTATGAGAAAGCAATTGTGAGGAAGGCAATTCCTTGAGCATCACTTCTGTCTGGGGAC
GTGGGTTAAGGCATAGCTGATCCTCTCTGGGACCAGGAAGAGAAATTAAGCTTAACAAGG
AGATGGTGGGTCATAGACTTCTCCTGAGTCTTAATTCATCTGCCATCTCATGTTGTGGGG
GAAGAGACAGTGAGATTCAGAGCTGGAATCTCCTAATATAATTGTGACAGGATTTGAAAA
AAAAATACTTTAATCCCAAGGGATCCAGGAAATAACCAAACCTGTTGTGAGAATAGGAAA

115600 AGAGAATTTATTTTGAAGAGATTCTCATGCAGAATCTAGGTGCTATAGAGGACGTACACC
TACTTTGAGAGTATGCTTGCATGAGTGGAAACCAATCATAAACAACATTCAACTTCATGA
GCAGATATGAAAGCATTTTCAGCATATCTAGCAATACTATAACTCTTTGTGCAAGCAGAG
TGGCCTACACAAGACAGTTTCAATATATTTTAAAAGAACGTCTTACATTTCATCAGTCCT
TTGAACACAGAAAAAAATGTTAAGGCCACTTAAGAGGCAAAACATCTTACAGAGTTCATT
[G,T]
ATATTCAAAGTCACCTACAGGCTACATCTTGGGTTCAGGAAGGGGCGGTGTACATAGTAA
GGACATACGCCTTCTGGGAGCCTTAAACAAACAAAAAAAATGTAGGTAACTCCTACATTT
TTCTTTTGTGGAAAAAACACAGTTACTCCAGCTTCCTTGGCTTTTTGCTTCTTTTTTATA
CCAACAAAATAAGGGCTATCCTCAACCCTCTGTTCTTCATTCTTCTCCCAGGGTATTGAT
TTCATAACATTGGGTTTTTCTTCTCTACTTCACTCATCCTCTTGCCTGTGAAGGTATGTA

115668 GAGTATGCTTGCATGAGTGGAAACCAATCATAAACAACATTCAACTTCATGAGCAGATAT
GAAAGCATTTTCAGCATATCTAGCAATACTATAACTCTTTGTGCAAGCAGAGTGGCCTAC
ACAAGACAGTTTCAATATATTTTAAAAGAACGTCTTACATTTCATCAGTCCTTTGAACAC
AGAAAAAAATGTTAAGGCCACTTAAGAGGCAAAACATCTTACAGAGTTCATTGATATTCA
AAGTCACCTACAGGCTACATCTTGGGTTCAGGAAGGGGCGGTGTACATAGTAAGGACATA
[A,C]
GCCTTCTGGGAGCCTTAAACAAACAAAAAAAATGTAGGTAACTCCTACATTTTTCTTTTG
TGGAAAAAACACAGTTACTCCAGCTTCCTTGGCTTTTTGCTTCTTTTTTATACCAACAAA
ATAAGGGCTATCCTCAACCCTCTGTTCTTCATTCTTCTCCCAGGGTATTGATTTCATAAC
ATTGGGTTTTTCTTCTCTACTTCACTCATCCTCTTGCCTGTGAAGGTATGTAAGGCTTCT
TTGTTCCAACTCTTTCCTCCACCCGCCCCCCCTCACATAAATGCATAACAAAGATTGTGA

115745 ATCTAGCAATACTATAACTCTTTGTGCAAGCAGAGTGGCCTACACAAGACAGTTTCAATA
TATTTTAAAAGAACGTCTTACATTTCATCAGTCCTTTGAACACAGAAAAAAATGTTAAGG
CCACTTAAGAGGCAAAACATCTTACAGAGTTCATTGATATTCAAAGTCACCTACAGGCTA
CATCTTGGGTTCAGGAAGGGGCGGTGTACATAGTAAGGACATACGCCTTCTGGGAGCCTT
AAACAAACAAAAAAAATGTAGGTAACTCCTACATTTTTCTTTTGTGGAAAAAACACAGTT
[A,G]
CTCCAGCTTCCTTGGCTTTTTGCTTCTTTTTTATACCAACAAAATAAGGGCTATCCTCAA
CCCTCTGTTCTTCATTCTTCTCCCAGGGTATTGATTTCATAACATTGGGTTTTTCTTCTC
TACTTCACTCATCCTCTTGCCTGTGAAGGTATGTAAGGCTTCTTTGTTCCAACTCTTTCC
TCCACCCGCCCCCCCTCACATAAATGCATAACAAAGATTGTGATTTAATTTAAGTTTCTT

TCTACTTTTAACATATTTGCAAACATCAATAGAAGCTAAAATGGGAAAAAGGAAATGTTT

117230 AATAATACTGTCGCTGCTAAGATAGGCATTGTGATATGGTGCTTAAACCTGCAAGTAAAG
GAAAAGAGTATGGAATCTGTGTGTCTTTTTCTAAGGGCTTTTTCCCAGAGTAGCTTGCAG
TCTGGCTTCTAGGGTTGCTGGCCTATAGCCAGAACCCTAGATTCACCCAGATTTACCTTC
AGAATTAACTAATCAGAGACTCAAATTCAATAGACTAAATGAAGTCAGGCTGCTAGAGGA
TGTCTGCTGACTTGGACATATGCAGAAAGACATGGATCCTTGAGAAAACATTGTTTCCAA
[A,C]
AGTGGCCACCAGCACTAGAGGAAGGACAGCACCACGGACAGCTCCCAGACATTTTAGGAT
TGCCTTCTGTGTTTGGTGCCCGAACACTGAGCAAAACAGCGAACTCAGGAAGTCTCCACA
CACTCTCATACCATCTTCATGCAGTCCAACTAAGAAAATTCTTACATAAAAATATAAGGCT
GTCTGCTTGGTAATTTAAACCCTTGGCTTATAGTCTTTTCAGTGAATTTCTTTCCTTGCA
AACTCGAGAGTTGGAGTCTCACGACTGCCCTTGCTTCACCAATTCCCCAGCTAGAGACAA

118908 CCCATTTGTAACATGAAACAAATAGTGCTGACCATTTGTATGCTAGGAATATTGTTAGGA
AACATAATATAGAATGTGAAATAAGTGGACTAGAAAGTCCTGAGATGTATTATCATTATT
GTTTAACTGTGTTTTTAAAGCAAAAATATTAAAACTCACTACTACAGGGCAAGATATATT
AACATCATTATTATTATTCATTATTGTATTATTCTAAATAGCCAATTTCAAAAGTCACAA
CCAGGCCAGGCAGTGAGGGACTCACGCCTGTAATCTCAGCACTTTGAGAGGCCGAGATGG
[A,G]
AGGGTCACTTATACCTAGGAATTTGAGACCAGCCTGGGCAACATAGGGAGACTCCATCTC
TATAAAAAATAAAACAAAATAAAAATCAGCTCAGTGTGGTTGTACATGCCTGTGGTCCCA
GCTACTCAGGAGGCTGAGGTGGGAGGATGGCTTGAGCCCAGGAGGTTGAGGTTGCAATGA
GCCATGATTGCACCACTGCACTCCAGCCTGGGTGACAAAGTGAGACCCTGTCTCAAACAA
AACAAAACAAAAAGATTACAACCAAAAACAAAGGGAAATAGAAGGATTGCCTCAAAAGAG

120430 CCCCAAGGCCCCAGAGCCAGAGCAAAAAGACAGCCAGGAAGAGAGGTTTGCTGCTGCTGC
TGCTGCTGCTACCCCACTTTTCTCATCACCTGCTTTAGATCTTTCTAGCTCCCCCTCTGA
TGACCTGACTGTGCCCCTCAAGACAATAAACGGAATGTAGGCCACATCATCTACCCTGCT
CCTTTTACAAAGGAGGGGACTGAGGTTCAGAAATAAGAGATGATTTACCCCAGCTTACAG
ATTTTCTTCATGGCAAAGCTGGAATGAGAACCCAAGTGTTCTGACTCCTGTTCTTTCAAA
[C,A]
CCCAGCTTCTACCGGTTATGCCAAAACATGACAGAAGTTGCCGTTGGCAAGGCACAGGCA
TGCCTCAGCATACCCTCCCCTCCAGGGCTGCTGAGTGGGCAACTCTGCCCACATTTCCTG
GCAAGGACAATCAAGGCCCATCCTGCTTTTTCCCATGAGATGTTTGGAGGAGGGCACTGG
CTCTGCAGTATATTCTCGTGATCTGGAATGACAGCCATCCCTCAGGGGACAGATAATGAC
CAGAACCACAATGGTTATTGCAGCAGTCAGGTCAGAAAATTTGAGAGGAGCCCTGCTGGC

120830 CAACTCTGCCCACATTTCCTGGCAAGGACAATCAAGGCCCATCCTGCTTTTTCCCATGAG
ATGTTTGGAGGAGGGCACTGGCTCTGCAGTATATTCTCGTGATCTGGAATGACAGCCATC
CCTCAGGGGACAGATAATGACCAGAACCACAATGGTTATTGCAGCAGTCAGGTCAGAAAA
TTTGAGAGGAGCCCTGCTGGCATCCAGTGAAGAGTGGCCACACCGAACTGATTTCACTTC
TCTCCTTAGACAACAAAATGCAGCCTGTGCATTCTCCTTTCTTTTTTTTTTTAATTATAC
[A,T]
TTAAGTTCTGGGGTACATGTGCAGAACATAGAGTTTTGTTACATAGGTATACACGTGCCA
TGGCGGTTTGCTGCACCCATCAACCCGTCATCTACATTAGGTATTTCTCCTAATGCTATC
CCTCCCCTATCCCTCACCCCTGACAGGCTCCAGTGTGTGATGTTCCTCTCCCTGTGTCCA
TGTGTTCTCATTGTTCAACTCCCACTTATGAGTGAGAACATGCAGTGTTTGGTTTTCTGT
TCTTGTGTTAGTTTGCTGAGAATGATGGTTTGCATCCTCCTTTCTTTCTGCTCCACTGTC

121926 TTGGTCTCAAAGATTCAGTCACAGCTGTTGTTTTCGTGGCATTTGGCACCTCTGTCCCAG
GTGAGAGTGAGAGGTGCTTGAATTTGCAAAGAGGATTTTACCTGGTTCAAATGACCCCTG
GACTCCATCTCATTATCTTCCACACCATCTCAGATCTGAACTTAACAGAGCCTCTGCCCT
TAAAGTGCACAAAAGTCAATCAAAGAGATGAATAATGACATTAGTAATGACAGCTAATAT
TTCTTGAGCACTTTCAATGTGACAGACACCATGTGTGTTCAGCAATTTACACATTTACAT
[T,C]
TTCCCCCTGTAATGTTTCCCAAAGCCCTATTAAATAGGGTAAGTTATTATCCCCACTTCA
CAGACAAAGAAACTGAGGCCCACAGAGGTTAAGCTACATGCCCAAGTAAGTGGTCCAATT
TCTTAACCTCCACATTATGTGAGTAGACCACAAACAGTGAAATTAAAAGAATGTAGATAT
TGTTCTCCTTCTATTTACCTCTGGCGATCTCTGAGAGGTTAAAGATTAGCCAGCTCAAAG
ATATCAAAGGAGAAATGCCCACATACATTCTTGGCCTCCTCTACTTGGAAGGACACTGTG

122102 CCCTTAAAGTGCACAAAAGTCAATCAAAGAGATGAATAATGACATTAGTAATGACAGCTA
ATATTTCTTGAGCACTTTCAATGTGACAGACACCATGTGTGTTCAGCAATTTACACATTT
ACATTTTTCCCCCTGTAATGTTTCCCAAAGCCCTATTAAATAGGGTAAGTTATTATCCCCA
CTTCACAGACAAAGAAACTGAGGCCCACAGAGGTTAAGCTACATGCCCAAGTAAGTGGTC
CAATTTCTTAACCTCCACATTATGTGAGTAGACCACAAACAGTGAAATTAAAAGAATGTA
[G,C]
ATATTGTTCTCCTTCTATTTACCTCTGGCGATCTCTGAGAGGTTAAAGATTAGCCAGCTC
AAAGATATCAAAGGAGAAATGCCCACATACATTCTTGGCCTCCTCTACTTGGAAGGACAC
TGTGAGTACAAAGTATCTCCTAGCAGGACAGCCAAAGGAAGTTCCACAGCTTTTATCTTT
TTATAGGATGAATTACATACTCTTTCTTTTTCTTAGGAACACTCAGAGACAAACAGAAAG
GAGCGGACATTCCTTTACTCATTGAACAAATATTTACTGAGCACCTATTATGCCTGTTAC

122950 GGGGCCTCTGAACTCTGAACTTCTGCCAGGGAGCTGGCATCCCAGTTGCCCCAGAAAGAA
ACAGAGCACATCCTCCTGCAGGGAAGTTAGGCTGAATCTCATCAGACAGGACTTTTCTGG
CTGGGCCAAGGGAAATCTTTCCTGTACCAAGCAAACATATCCTTCAAGAGAGTAGCTGAA

TTCACATCAAATTCTAGGAAAACCTCTTTCCAAAACCCCAGCGCAGGCCAGCGGTATTAT
TTGTCCATTAGTGATGCAAGAGATTTAGCTATCGTGGAAATGCATCAGAAGGTTGGAAAT
[T,C]
AGATGGATGATCCCAGGAAGGCCTGTGGATGAGATGCCCTGTGATCTCTGTTCTCCAAGC
CTTGGGGGACCTGAACTATCAGAGGGGAGGGAGGAAATATGGGGGAAAGCATAGAGGTGG
GAAGAAATATCAGAGGATCAGAAGCAAAAAACAACAATAACAACAGAAACAAAAACAAAC
AAACAAACAAAAAAACAAGGCCATAGGCAAGAAAGGGTAAGAGGTTTTCTCTGGGAGATC
TAAAAAAAATGGCAATAATGAGGTAAGCCAGGCAGATACCTTTGGGCATCTCCAAGTCCT

123366 GGTGGGAAGAAATATCAGAGGATCAGAAGCAAAAAACAACAATAACAACAGAAACAAAAA
CAAACAAACAAACAAAAAAACAAGGCCATAGGCAAGAAAGGGTAAGAGGTTTTCTCTGGG
AGATCTAAAAAAAATGGCAATAATGAGGTAAGCCAGGCAGATACCTTTGGGCATCTCCAA
GTCCTTGCAATTGGCCAAGACAACAGCTAACAACATTTGAGGCTTTAAGAAGGTTACCCT
GTGATCCACTCATCTGATTTAGTGGCTTTGGCTGAAGCTCTTTGGATATAGTTGAAGGTA
[C,T]
GGAAAGGGTCCTTACATGAGGACTTTAGGGTCAAGTCTCTTGCTAACATCCTATGTGACC
TTGGGTAAATTCTTTGACCCTTATTTTTCTTACCTGTAAAATAAAAGAATTGGGCTAGAT
GTCTCTGACAGTCCTCCCTGTATCTACAATCTGTGCCAAGATCTAAAGTCAAACACCCTG
CAAGGCCCTGTGATACATATATAAACCACAAAGACAGAGCCCCGTCTTCCTTGAGTCCAC
AGTTCACCCTGCATGTCCCCATCATGGTTCCCCAACATGTCCTCTGTCCCCAAAATCCAG

124947 GGCCCTGGCAATTAACTGAAAGGGCAAAGTCTTAATCAATCAAACAATGGAGGAATCACC
GACTTTACACAGTATTTAATTGAATACAAACAAGCAACAGCAACAAATCCACCTCCACCC
CATCTCCCCCTCATATCCCTGACCCAAAGCAAAGGTCAGAGCCTTTCGCCTCCTTCTATT
CCATCTTTTGATTATTCCTTTGCCTCTCATTTCTTTGGAAGCAGGGTTTCTCCTCTCTGC
CCAATTCCATATGTCCCTATTATCTCACTCAGCTGACAAGACGTGAAAATGAGTCACATT
[C,T]
ATGTGGCTGGGGTGGGGTTCTTTTTTCATTGTAATCATTATTGTGGTTGCTTTCGTTTTG
CCGTTAGGTTTTGCTTATTATTTTGTTTTGTCTTTTTTTTCTGAAGTGAGTGAAAAAGGT
GCCACAAAGGAATTCCAGGTCCGAGCCAACAGAGAGAAACATGAATTTTTAGACACATGC
TCTCCTGCCACCTCTTGGCTCCATCAAGATCCAGTTCCCCATCTCACTGTTTTCTCTGAG
TTCTTGGGAGGAGTGATGGTGTTGGGGTAGAAATAAGCTCACTCACCCACGCAGGGTACT

125010 TTTACACAGTATTTAATTGAATACAAACAAGCAACAGCAACAAATCCACCTCCACCCCAT
CTCCCCCTCATATCCCTGACCCAAAGCAAAGGTCAGAGCCTTTCGCCTCCTTCTATTCCA
TCTTTTGATTATTCCTTTGCCTCTCATTTCTTTGGAAGCAGGGTTTCTCCTCTCTGCCCA
ATTCCATATGTCCCTATTATCTCACTCAGCTGACAAGACGTGAAAATGAGTCACATTCAT
GTGGCTGGGGTGGGGTTCTTTTTTCATTGTAATCATTATTGTGGTTGCTTTCGTTTTGCC
[A,G]
TTAGGTTTTGCTTATTATTTTGTTTTGTCTTTTTTTTCTGAAGTGAGTGAAAAAGGTGCC
ACAAAGGAATTCCAGGTCCGAGCCAACAGAGAGAAACATGAATTTTTAGACACATGCTCT
CCTGCCACCTCTTGGCTCCATCAAGATCCAGTTCCCCATCTCACTGTTTTCTCTGAGTTC
TTGGGAGGAGTGATGGTGTTGGGGTAGAAATAAGCTCACTCACCCACGCAGGGTACTAAA
GATCTTACAGGAGCTTCAACTGGAGCAGGAGGAGCTTTTTATGCTTATGTTGAATCAAGT

126043 AAAGCATTTTTACAAGATAGGAACTGGAATTCCTCATTTCTCCCATGTTCCTGCTTGTTC
TTAAACTTCATGAAGCTATTTTTCCAGCCTATGGGGTAGTTCTTGCTCCAGTAAGAGGAA
TCTTAGTTGTCATAATCCCTTGGAGCCTGGGTTTTTGGAGAAAGAGATCTCCGTGCCCTA
CAGACCTTTTCTCAACGAATGTGGGAAGGACCTGGCTTTAAAACACGCACACAAACACAC
AAATAAACAGACATAAGATGTCATCACGAAACTGCCCACGGATCTTTAGGCTTTCTGCAT
[T,C]
GACATAAATACATTTTCTAAGGGGGGGGGGGGAAGAAATTAAAAAACACCTGTTAATTTTA
AACACATTTTTTAAGAAAAAAATAATTAAAAAAGAAACAGTGCTCATGTCATAAGCTATG
TTGACAGTTGCCAGTGGAAATGTTGGGTTGGTTCAAAAAAAAAATAAAAGCTATACTATA
TCTCTCTACATACAGCTTGCTTCTACCTGTGTTTCTTCAGTGAAAGGTCCAGGGGGCCAC
TGTGGGCTTCTTGTGAGGAGACGTGACTCAGGTGAAGGTGTCACCTCCTCTCACACTCAG

126064 AACTGGAATTCCTCATTTCTCCCATGTTCCTGCTTGTTCTTAAACTTCATGAAGCTATTT
TTCCAGCCTATGGGGTAGTTCTTGCTCCAGTAAGAGGAATCTTAGTTGTCATAATCCCTT
GGAGCCTGGGTTTTTGGAGAAAGAGATCTCCGTGCCCTACAGACCTTTTCTCAACGAATG
TGGGAAGGACCTGGCTTTAAAACACGCACACAAACACACAAATAAACAGACATAAGATGT
CATCACGAAACTGCCCACGGATCTTTAGGCTTTCTGCATTGACATAAATACATTTTCTAA
[-,G]
GGGGGGGGGGAAGAAATTAAAAAACACCTGTTAATTTTAAACACATTTTTTAAGAAAAAA
ATAATTAAAAAAGAAACAGTGCTCATGTCATAAGCTATGTTGACAGTTGCCAGTGGAAAT
GTTGGGTTGGTTCAAAAAAAAAATAAAAGCTATACTATATCTCTCTACATACAGCTTGCT
TCTACCTGTGTTTCTTCAGTGAAAGGTCCAGGGGGCCACTGTGGGCTTCTTGTGAGGAGA
CGTGACTCAGGTGAAGGTGTCACCTCCTCTCACACTCAGGTGCCAATGTGTCAGACCCAG

126283 AAATAAACAGACATAAGATGTCATCACGAAACTGCCCACGGATCTTTAGGCTTTCTGCAT
TGACATAAATACATTTTCTAAGGGGGGGGGGGGAAGAAATTAAAAAACACCTGTTAATTTT
AAACACATTTTTTAAGAAAAAAATAATTAAAAAAGAAACAGTGCTCATGTCATAAGCTAT
GTTGACAGTTGCCAGTGGAAATGTTGGGTTGGTTCAAAAAAAAAATAAAAGCTATACTAT
ATCTCTCTACATACAGCTTGCTTCTACCTGTGTTTCTTCAGTGAAAGGTCCAGGGGGCCA
[C,G]
TGTGGGCTTCTTGTGAGGAGACGTGACTCAGGTGAAGGTGTCACCTCCTCTCACACTCAG
GTGCCAATGTGTCAGACCCAGTATATTCTAAGCAAAAATACTTCAGGAAAATGCCACTTG

TCAAAACCTGGACTTTGCGAAGTTGGAAGATGTAAGTAGTAGTAAAAGCTGTGGTAATTA
TGGAGGAAGGAGGTTTCTGTATCAGAAAGGCATTGGCCGTGACAGACTC

Chromosome map:
Chromosome 14

SEQUENCE LISTING

<110> PE CORPORATION (NY)

<120> ISOLATED HUMAN TRANSPORTER PROTEINS,
NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
AND USES THEREOF

<130> CL000891PCT

<140> TO BE ASSIGNED
<141> 2001-17-10

<150> 60/240,836
<151> 2000-17-10

<150> 09/804,474
<151> 2001-13-03

<160> 4

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 2782
<212> DNA
<213> Human

<400> 1
gtctcgtgta tggcgtggtt aaggttgcag cctctcacct ctgccttcct ccattttggg 60
ctggttacct ttgtgctctt cctgaatggt cttcgagcag aggctggtgg ctcaggggac 120
gtgccaagca cagggcagaa caatgagtcc tgttcagggt catcggactg caaggagggt 180
gtcatcctgc caatctggta cccggagaac ccttcccttg ggacaagat tgccagggtc 240
attgtctatt ttgtggccct gatatacatg ttccttgggg tgtccatcat tgctgaccgc 300
ttcatggcat ctattgaagt catcacctct caagagaggg aggtgacaat taagaaaccc 360
aatggagaaa ccagcacaac cactattcgg gtctggaatg aaactgtctc caacctgacc 420
cttatggccc tgggttcctc tgctcctgag atactcctct ctttaattga ggtgtgtggt 480
catgggttca ttgctggtga tctgggacct tctaccattg tagggagtgc agccttcaac 540
atgttcatca tcattggcat ctgtgtctac gtgatcccag acggagagac tcgcaagatc 600
aagcatctac gagtcttctt catcaccgct gcttggagta tctttgccta catctggctc 660
tatatgattc tggcagtctt ctcccctggt gtggtccagg tttgggaagg cctcctcact 720
ctcttcttct ttccagtgtg tgtccttctg gcctgggtgg cagataaacg actgctcttc 780
tacaaataca tgcacaaaaa gtaccgcaca gacaaacacc gaggaattat catagagaca 840
gagggtgacc accctaaggg cattgagatg gatgggaaaa tgatgaattc ccattttcta 900
gatgggaacc tggtgcccct ggaagggaag gaagtggatg agtcccgcag agagatgatc 960
cggatcctca aggatctgaa gcaaaaacac ccagagaagg acttagatca gctggtggag
1020
atggccaatt actatgctct ttcccaccaa cagaagagcc gcgccttcta ccgtatccaa
1080
gccactcgta tgatgactgg tgcaggcaat atcctgaaga aacatgcagc agaacaagcc
1140
aagaaggcct ccagcatgag cgaggtgcac accgatgagc ctgaggactt tatttccaag
1200
gtcttctttg acccatgttc ttaccagtgc ctggagaact gtggggctgt actcctgaca
1260
gtggtgagga aagggggaga catgtcaaag accatgtatg tggactacaa aacagaggat
1320
ggttctgcca atgcaggggc tgactatgag ttcacagagg gcacggtggt tctgaagcca
1380

ggagagaccc agaaggagtt ctccgtgggc ataattgatg acgacatttt tgaggaggat
1440
gaacacttct ttgtaaggtt gagcaatgtc cgcatagagg aggagcagcc agaggagggg
1500
atgcctccag caatattcaa cagtcttccc ttgcctcggg ctgtcctagc ctccccttgt
1560
gtggccacag ttaccatctt ggatgatgac catgcaggca tcttcacttt tgaatgtgat
1620
actattcatg tcagtgagag tattggtgtt atggaggtca aggttctgcg gacatcaggt
1680
gcccggggta cagtcatcgt cccctttagg acagtagaag ggacagccaa gggtggcggt
1740
gaggactttg aagacacata tggggagttg gaattcaaga atgatgaaac tgtgaaaacc
1800
ataagggtta aaatagtaga tgaggaggaa tacgaaaggc aagagaattt cttcattgcc
1860
cttggtgaac cgaaatggat ggaacgtgga atatcagatg tgacagacag gaagctgact
1920
atggaagaag aggaggccaa gaggatagca gagatgggaa agccagtatt gggtgaacac
1980
cccaaactgg aagtcatcat tgaagagtcc tatgagttca agactacggt ggacaaactg
2040
atcaagaaga caaacctggc cttggttgtg gggacccatt cctggaggga ccagttcatg
2100
gaggccatca ccgtcagtgc agcaggggat gaggatgagg atgaatccgg ggaggagagg
2160
ctgccctcct gctttgacta cgtcatgcac ttcctgactg tcttctggaa ggtgctgttt
2220
gcctgtgtgc cccccacaga gtactgccac ggctgggcct gcttcgccgt ctccatcctc
2280
atcattggca tgctcaccgc catcattggg gacctggcct cgcacttcgg ctgcaccatt
2340
ggtctcaaag attcggtcac agctgttgtt ttcgtggcat ttggcacctc tgtcccagat
2400
acgtttgcca gcaaagctgc tgccctccag gatgtatatg cagacgcctc cattggcaac
2460
gtgacgggca gcaacgccgt caatgtcttc ctgggcatcg gcctggcctg gtccgtggcc
2520
gccatctact gggctctgca gggacaggag ttccacgtgt cggccggcac actggccttc
2580
tccgtcaccc tcttcaccat ctttgcattt gtctgcatca gcgtgctctt gtaccgaagg
2640
cggccgcacc tgggaggggа gcttggtggc ccccgtggct gcaagctcgc cacaacatgg
2700
ctctttgtga gcctgtggct cctctacata ctctttgcca cactagaggc ctattgctac
2760
atcaaggggt tctaagccac ac
2782

<210> 2
<211> 921
<212> PRT
<213> Human

<400> 2
Met Ala Trp Leu Arg Leu Gln Pro Leu Thr Ser Ala Phe Leu His Phe
1 5 10 15
Gly Leu Val Thr Phe Val Leu Phe Leu Asn Gly Leu Arg Ala Glu Ala
20 25 30
Gly Gly Ser Gly Asp Val Pro Ser Thr Gly Gln Asn Asn Glu Ser Cys
35 40 45

Ser Gly Ser Ser Asp Cys Lys Glu Gly Val Ile Leu Pro Ile Trp Tyr
    50   55   60
Pro Glu Asn Pro Ser Leu Gly Asp Lys Ile Ala Arg Val Ile Val Tyr
65   70   75   80
Phe Val Ala Leu Ile Tyr Met Phe Leu Gly Val Ser Ile Ile Ala Asp
                85   90   95
Arg Phe Met Ala Ser Ile Glu Val Ile Thr Ser Gln Glu Arg Glu Val
            100   105   110
Thr Ile Lys Lys Pro Asn Gly Glu Thr Ser Thr Thr Thr Ile Arg Val
        115   120   125
Trp Asn Glu Thr Val Ser Asn Leu Thr Leu Met Ala Leu Gly Ser Ser
    130   135   140
Ala Pro Glu Ile Leu Leu Ser Leu Ile Glu Val Cys Gly His Gly Phe
145   150   155   160
Ile Ala Gly Asp Leu Gly Pro Ser Thr Ile Val Gly Ser Ala Ala Phe
                165   170   175
Asn Met Phe Ile Ile Ile Gly Ile Cys Val Tyr Val Ile Pro Asp Gly
            180   185   190
Glu Thr Arg Lys Ile Lys His Leu Arg Val Phe Phe Ile Thr Ala Ala
        195   200   205
Trp Ser Ile Phe Ala Tyr Ile Trp Leu Tyr Met Ile Leu Ala Val Phe
    210   215   220
Ser Pro Gly Val Val Gln Val Trp Glu Gly Leu Leu Thr Leu Phe Phe
225   230   235   240
Phe Pro Val Cys Val Leu Leu Ala Trp Val Ala Asp Lys Arg Leu Leu
                245   250   255
Phe Tyr Lys Tyr Met His Lys Lys Tyr Arg Thr Asp Lys His Arg Gly
            260   265   270
Ile Ile Ile Glu Thr Glu Gly Asp His Pro Lys Gly Ile Glu Met Asp
        275   280   285
Gly Lys Met Met Asn Ser His Phe Leu Asp Gly Asn Leu Val Pro Leu
    290   295   300
Glu Gly Lys Glu Val Asp Glu Ser Arg Arg Glu Met Ile Arg Ile Leu
305   310   315   320
Lys Asp Leu Lys Gln Lys His Pro Glu Lys Asp Leu Asp Gln Leu Val
                325   330   335
Glu Met Ala Asn Tyr Tyr Ala Leu Ser His Gln Gln Lys Ser Arg Ala
            340   345   350
Phe Tyr Arg Ile Gln Ala Thr Arg Met Met Thr Gly Ala Gly Asn Ile
        355   360   365
Leu Lys Lys His Ala Ala Glu Gln Ala Lys Lys Ala Ser Ser Met Ser
    370   375   380
Glu Val His Thr Asp Glu Pro Glu Asp Phe Ile Ser Lys Val Phe Phe
385   390   395   400
Asp Pro Cys Ser Tyr Gln Cys Leu Glu Asn Cys Gly Ala Val Leu Leu
                405   410   415
Thr Val Val Arg Lys Gly Gly Asp Met Ser Lys Thr Met Tyr Val Asp
            420   425   430
Tyr Lys Thr Glu Asp Gly Ser Ala Asn Ala Gly Ala Asp Tyr Glu Phe
        435   440   445
Thr Glu Gly Thr Val Val Leu Lys Pro Gly Glu Thr Gln Lys Glu Phe
    450   455   460
Ser Val Gly Ile Ile Asp Asp Asp Ile Phe Glu Glu Asp Glu His Phe
465   470   475   480
Phe Val Arg Leu Ser Asn Val Arg Ile Glu Glu Glu Gln Pro Glu Glu
                485   490   495
Gly Met Pro Pro Ala Ile Phe Asn Ser Leu Pro Leu Pro Arg Ala Val
            500   505   510
Leu Ala Ser Pro Cys Val Ala Thr Val Thr Ile Leu Asp Asp Asp His
        515   520   525
Ala Gly Ile Phe Thr Phe Glu Cys Asp Thr Ile His Val Ser Glu Ser

```
    530                 535                 540
Ile Gly Val Met Glu Val Lys Val Leu Arg Thr Ser Gly Ala Arg Gly
545                 550                 555                 560
Thr Val Ile Val Pro Phe Arg Thr Val Glu Gly Thr Ala Lys Gly Gly
                565                 570                 575
Gly Glu Asp Phe Glu Asp Thr Tyr Gly Glu Leu Glu Phe Lys Asn Asp
            580                 585                 590
Glu Thr Val Lys Thr Ile Arg Val Lys Ile Val Asp Glu Glu Glu Tyr
        595                 600                 605
Glu Arg Gln Glu Asn Phe Phe Ile Ala Leu Gly Glu Pro Lys Trp Met
    610                 615                 620
Glu Arg Gly Ile Ser Asp Val Thr Asp Arg Lys Leu Thr Met Glu Glu
625                 630                 635                 640
Glu Glu Ala Lys Arg Ile Ala Glu Met Gly Lys Pro Val Leu Gly Glu
                645                 650                 655
His Pro Lys Leu Glu Val Ile Ile Glu Glu Ser Tyr Glu Phe Lys Thr
            660                 665                 670
Thr Val Asp Lys Leu Ile Lys Lys Thr Asn Leu Ala Leu Val Val Gly
        675                 680                 685
Thr His Ser Trp Arg Asp Gln Phe Met Glu Ala Ile Thr Val Ser Ala
    690                 695                 700
Ala Gly Asp Glu Asp Glu Asp Glu Ser Gly Glu Glu Arg Leu Pro Ser
705                 710                 715                 720
Cys Phe Asp Tyr Val Met His Phe Leu Thr Val Phe Trp Lys Val Leu
                725                 730                 735
Phe Ala Cys Val Pro Pro Thr Glu Tyr Cys His Gly Trp Ala Cys Phe
            740                 745                 750
Ala Val Ser Ile Leu Ile Ile Gly Met Leu Thr Ala Ile Ile Gly Asp
        755                 760                 765
Leu Ala Ser His Phe Gly Cys Thr Ile Gly Leu Lys Asp Ser Val Thr
    770                 775                 780
Ala Val Val Phe Val Ala Phe Gly Thr Ser Val Pro Asp Thr Phe Ala
785                 790                 795                 800
Ser Lys Ala Ala Ala Leu Gln Asp Val Tyr Ala Asp Ala Ser Ile Gly
                805                 810                 815
Asn Val Thr Gly Ser Asn Ala Val Asn Val Phe Leu Gly Ile Gly Leu
            820                 825                 830
Ala Trp Ser Val Ala Ala Ile Tyr Trp Ala Leu Gln Gly Gln Glu Phe
        835                 840                 845
His Val Ser Ala Gly Thr Leu Ala Phe Ser Val Thr Leu Phe Thr Ile
    850                 855                 860
Phe Ala Phe Val Cys Ile Ser Val Leu Leu Tyr Arg Arg Arg Pro His
865                 870                 875                 880
Leu Gly Gly Glu Leu Gly Gly Pro Arg Gly Cys Lys Leu Ala Thr Thr
                885                 890                 895
Trp Leu Phe Val Ser Leu Trp Leu Leu Tyr Ile Leu Phe Ala Thr Leu
            900                 905                 910
Glu Ala Tyr Cys Tyr Ile Lys Gly Phe
        915                 920


<210> 3
<211> 126512
<212> DNA
<213> Human

<220>
<221> misc_feature
<222> (1)...(126512)
<223> n = A,T,C or G
```

<400> 3
ttggatgaga tctaaagcat tattaagagt ggggagtgca aagaagaaac cctcatttca 60
aagatgaatg agaataatgg catgtacaaa ggtcctgggg tggacagtca cttggtataa 120
tccaagagtg aacctgaagg ctattgttgt tgaaatgtaa taagggagag agtgacggga 180
tgaaggggga tgagtgggaa gcagtgaatt cctgcaaggc tttgaaggtc atgggaaaga 240
atttggtctt tatatcaaga gcaagagaag actactaaag ggcttcaaac aggggagcga 300
tatgcttaag tctgtttgtt tgttttttta aaaaaagatt acggtggcta tatgaggaaa 360
gtggaattga gaactagcga gagttggagt ggtgagctcc attaggaggc tactgaagta 420
gattcatgag gtaaggagtg atggtggcct gggctgggat gatggtggta gaaatggaga 480
aagagttgat aggatttagt gattggataa gggacagaag agagatgaag gctttcagac 540
taacatctgc tttctaacat gagtaactgg gtggctgaag atgctatttt ctgagctggg 600
aaacaggaga aaaaggagca aatatggggg atgaagactt tgagtcttta aggtgctgta 660
caaacacaaa tcagcattcc tttattacta agggtatccc acacagttgt agcagaggga 720
gaaagatcgc cccccccca cttttttttt ttttttagct attccatggt attttcattc 780
tcatcccacc caaatgaggc agtgagtggt aagatgagta tataatagtt tcaattgcat 840
ttcatcccat tcttctgagc tcaagctcac cttttagtgg tttgaggcca gtagatgaag 900
ctgcatatca cccccaaaat cttgtctcta gtttaacaaa acttatttga gagacatttg 960
catgttttat taataatgat ttttaccact tgttcctttc catgtttggg tttgaaattt
1020
gagtggctgg cggatgatca tcttcctgtt actgcctgct taaactgctc ataagcaggt
1080
tttactggag ggctcagagc tgctgtgaac ttggtcttgg gcacaactta catggcctct
1140
gtttggctat ggggtgggtg gcattcacca tttatcaact cttttgattt cccaagctat
1200
ctcagaatta tagcttgcct ccagaagtct tgcattcggg gaggaagttt ctttccaagg
1260
gagctcagtt ttcaaggttt attgctctgt ttaatggatg agatctaaag cattattaag
1320
agtggggagt gcaaagaaga aacactcatt tcaaaatcga ttgagaataa tggcatgtac
1380
aaaggtcctg gggtggacag tcacttggta taatcctgga gtgaacatga aggccaagga
1440
aatatgtata cattaaacag agcaaggttt tcaattttct ggggactagt ccatgaaaat
1500
tcaattcaat atactctctt gcaaacctat gttatccaag atactcaagt ataatgacaa
1560
cagggtaagg aagtccgaac accccagaaa cagtataaat gggcatgaag attcaggtta
1620
tacatggcct attttaagtt gcttcttgag aactctcaca ggtaatacca gtttgggaga
1680
caggacttga aggctattgc tgcatttcca tccccagtat tcccagctat ttcaagccat
1740
ttttcaacgg agtctccacc agatggtttg gaggacagag cagctatttg tgcctcccat
1800
tgacatctat ttttccaagt gagagactgc cccatatgtt agtgcaatat gtcactggag
1860
gtgaagcatc agttgtattg gtgggaacct gccgtttgct gtcccctttt tcctcatgcc
1920
ttttcctgcc tctctgatct tttctaggtc tctggcctat caggaggaca actggtgctg
1980
caatagaagc cagtggctaa gtctcgtgta tggcgtggtt aaggttgcag cctctcacct
2040
ctgccttcct ccattttggg ctggttacct ttgtgctctt cctgaatggt cttcgagcag
2100
aggctggtgg ctcaggggac gtgccaagca cagggcagaa caatgagtcc tgttcagggt
2160
catcggactg caaggagggt gtcatcctgc caatctggta cccggagaac ccttcccttg
2220
gggacaagat tgccagggtc attgtctatt ttgtggccct gatatacatg ttccttgggg
2280

tgtccatcat tgctgaccgc ttcatggcat ctattgaagt catcacctct caagagaggg 2340
aggtgacaat taagaaaccc aatggagaaa ccagcacaac aactattcgg gtctggaatg 2400
aaactgtctc caacctgacc cttatggccc tgggttcctc tgctcctgag atactcctct 2460
ctttaattga ggtgtgtggt catgggttca ttgctggtga tctgggacct tctaccattg 2520
tagggagtgc agccttcaac atgttcatca tcattggcat ctgtgtctac gtgatcccag 2580
acggagagac tcgcaagatc aagcatctac gagtcttctt catcaccgct gcttggagta 2640
tctttgccta catctggctc tatatgattc tggcagtctt ctcccctggt gtggtccagg 2700
tttgggaagg cctcctcact ctcttcttct ttccagtgtg tgtccttctg gcctgggtgg 2760
cagataaacg actgctcttc tacaaataca tgcacaaaaa gtaccgcaca gacaaacacc 2820
gaggaattat catagagaca gagggtgacc accctaaggg cattgagatg gatgggaaaa 2880
tgatgaattc ccattttcta gatgggaacc tggtgcccct ggaagggaag gaagtggatg 2940
agtcccgcag agagatgatc cggattctca aggatctgaa gcaaaaacac ccagagaagg 3000
acttagatca gctggtggag atggccaatt actatgctct ttcccaccaa cagaagagcc 3060
gcgccttcta ccgtatccaa gccactcgta tgatgactgg tgcaggcaat atcctgaaga 3120
aacatgcagc agaacaagcc aagaaggcct ccagcatgag cgaggtgcac accgatgagc 3180
ctgaggactt tatttccaag gtcttctttg acccatgttc ttaccagtgc ctggagaact 3240
gtggggctgt actcctgaca gtggtgagga aaggggagga catgtcaaag accatgtatg 3300
tggactacaa aacagaggat ggttctgcca atgcaggggc tgactatgag ttcacagagg 3360
gcacggtggt tctgaagcca ggagagaccc agaaggagtt ctccgtgggc ataattgatg 3420
acgacatttt tgaggaggat gaacacttct ttgtaaggtt gagcaatgtc cgcatagagg 3480
aggagcagcc agaggagggg atgcctccag caatattcaa cagtcttccc ttgcctcggg 3540
ctgtcctagc ctccccttgt gtggccacag ttaccatctt ggatgatgac catgcaggca 3600
tcttcacttt tgaatgtgat actattcatg tcagtgagag tattggtgtt atggaggtca 3660
aggttctgcg gacatcaggt gcccggggta cagtcatcgt cccctttagg acagtagaag 3720
ggacagccaa gggtggcggt gaggactttg aagacacata tggggagttg gaattcaaga 3780
atgatgaaac tgtgtaagta accttcctgt attctgcccc tccctgaccc catcttttgc 3840
catctctttc tgtctttctg tactgcactt tacaacattt ccttgtgttt gtgttaatgt 3900
caaactttgg ttccatcaca ggtatgcagg atcagcagac accactggac aggttctgct 3960
tccaaactct tcttcagttt tctcacttta aattgtttct gggcaaggaa tcctgtgaca 4020
agagctaagg acacaaaaca ttttcttctc tgaaacacaa aatgatagct ggtggagctg 4080

tgggatgaca gaagttttgt gatatcagat tttggagaat tcttgtgact aagaaggact
4140
agagaactgc ttgggcctct tcttcctccc ttcctcatat gaagggtatc tatgagcttt
4200
gaaaccaatc ctttccattc tgggcagcaa tagcccatca gaacattcta aagaaaacaa
4260
gtggcattgg ctttgttccc tggtactata ttgccagtct cactgtgtaa ccagattcca
4320
ggcacgtctt ctttaatttg gaaattgcaa aattgataga aatttagcaa tctttttaaa
4380
tgaccataga ctatttaatg gtgtgaggct tgcccagcct agttgaattg agtcagtatg
4440
gtttggatac tggaaagtat cttggagaag cagagctccc agggcagtgg ctacttgtct
4500
ttagtcacag gtctaagctc caaaatctgg tgaagcagtg aaggagaaac atcctaggaa
4560
ttgtgggagg aaatatatct tctgtgtggt cctctctttt cacagtctag gactctcctg
4620
aagtacctct tcttgggcta ctgcccccatt cagcccttca gaaactgtgg gtattacact
4680
tctgtcacct ctattaccct aaggcctctg cccattgaac cctcttgcaa attggttatt
4740
ctgtcctttt tccagttgga tagctttaaa agggaaagca gaatgacttt cctcaggatt
4800
tgtagcttat gagaaagtag actttcttgg gtggcctaga aggttggaga agacaaacgg
4860
gaacttcctc tgaatgactg aacatatcca caaataataa gcgtggcagg agatggtgtg
4920
aagagtaaaa ggagcatata ggaagttgtg tgtgtggggt gtctgtttca agaacctgct
4980
aattatacct tcagtaagaa atgaagccat acaacctcta gaagaggagg aggaaggaac
5040
tcatggaaaa gtggggagcc atagaagcta gggagaggtg tcctaggagt gcttctgccc
5100
aggtccagcc atgagacaga gctcaaaaag agctgggcac tgctggtgac agaactgagt
5160
gacccggggg atcctgcatc tgttcttact caatcccttc ttaataatgt gacttggggc
5220
aggtcattta ttggttctgg aacttaactt tctgatatgc aaactgggaa taacaatact
5280
ttccttgcct ggaggcaagg tcagtccttt ttgcagttcc ttccagctct aagattttct
5340
gaaccataga cataagcact cagtgtaggt catattcgca cttgccaaaa atggatcagg
5400
gaatattgtc tcctgaaggg aaatggccat tgacaaattg atttattaga gctctgttta
5460
gtcattttgc tgggaaggat aatcatttgt taacgtaagt agaaacctgt gccttctgga
5520
gaatactatc catttatatg tactctgggg agagtgttta tacatacaaa tgaaggacag
5580
ggcttcactg ggaaaacaaa ctccatggaa tttcacatga ttatcgcgat gtcagtgtgg
5640
aagaagatat ggtaaggcat taaatgacat taagaccaca aaatttgcca taatttgacg
5700
gacttgtggt tcttctgatt cagaaccctt tctacccatg tcacggatag gtagtttttc
5760
agagatcaga ggcttagttc attctattaa tttcctcatt ctattaataa tcaattatgc
5820
acctagggtc tctgaatacg actaaacctt cctcaaactt atttgcattt tcagtttgta
5880

taatatcttg gtgcaaatga gcctcgcaaa tgatcacttc tgggtaatac tcattctaaa
5940
ggtatgtcaa ccttgagaat tctggtctag atattctagg gtttggtgaa caaatctatg
6000
ttcccatcca tccctttca tttatttttt agacttcatt cattgcagaa taatgagtcc
6060
aaaacctgct catctgttct cacgtggcac ccctattctt gatattttaa attgcaattt
6120
tacaactaga ggcagtatta cggagcagaa aaatcgtggg ttctaagtac tctgggttag
6180
gattctggct ccactactga tttaataatg tagtttgggg aaattttatt aacctatgaa
6240
attatttcct cattggcaaa atggggataa taatatctct cttgcagggc cattatgacg
6300
attcaaggta ttgtatgcgg tgtacctggt acacggtata tgctcaggaa acaagactct
6360
tcatagtaat attgacgaat taacaatatt cttcagaaga cactgtggag ttgtttaggt
6420
tacttggctc tttgtgtgac cctaagtaat gagcatgcca gtttggggtt actatgaaga
6480
gtacttacct aaactcataa aatattagag ctagaaagga ccttagaata tcttctgcag
6540
tcatggttct taaattttaa tgtgttgctc aatcatccag ggatctcact gaagggcaga
6600
ttaggatcca ggaggtctag gggagggatt gagattccgc atttctaaca agttctggat
6660
gctgcgggcc ccaacttaga ggtgaaaggt tctgaagctc ttgaccaaac caggagaccc
6720
agcaaagaag tggtttttca gacaacttgc ttaattgaat aatgattgtt tgctctttaa
6780
ttccaacttt caatgccaat ttagcaagaa ccagaggctg tgctaattgc cacaccagtc
6840
tggaaaccga aatggatagc ttcagggtac ttggacaaag ttggaacatc tgctttctaa
6900
tctctccctc tttgtatagc tttatttgcc taccaagcct ggtagtattg aaaatctgcc
6960
ctcactatac tcccctaaat ataatcaagt tgaggccagg cctgtgctct atcaataata
7020
taggatccac gaattcacat gtttggtttt atgctttact tcttcaaagg tgcttttagc
7080
agcatggaag aatggaaaag cacgagcttt ggaatatgaa agcagatgtg aatccatcac
7140
ttaccagtaa cttttaacaa gtcacatcac ttttctgagt accaggtttt tgttggacaa
7200
cagaaataat attctctatc cttcaaggga atactaaata taagtatgag aaaaatgcac
7260
agtgccttct cgtagatggt gttcagtcat tcaacaaaca tttgttagat atttgctatg
7320
tactagctac attactaggc actggggtta aataagtgaa taagacaagc tgacatttca
7380
gcgctcaagg atcttactgt caagtggaga ggatcaaagg gtacagacaa atcaaggaac
7440
gtgagagaag tggtatggct gagatggatt gaataaagga gcaatgagag ctcccctgcaa
7500
tgtgtgtggt accactgagg attctaaatt aaccttcatt aaggacttag tagtgacaga
7560
ggtgaagtgg ggataggtac atgattaatt tacatccata ttacaatgaa accttaacat
7620
ttaagaggga tattattgat gtcttcatga tccagaagaa tcctcacctt tgcaaccatc
7680

actatagtca cttcttgaga attatggcct ttaagactgt agcatgcaat gacaaaacct
7740
cacagaggta tgggttctgc ccgcacacta atttcactca ttaaacaagt gactggctcc
7800
tatatcccag gctctcagca cgcctttgca aaataacaga ttattgcagc tcttggacct
7860
ttgatgcctc tgggaatagt caaagccaca gatgtcaaat atgtaaatgc caagatctat
7920
tataattaaa tagtgcaggc ctccttcaaa gaaaaaaagc atgttggctg tgctgcacgt
7980
tctccaacca aatcagaatg ttaaagctcg aaggtatctg acctcccatt ttttaaatta
8040
tgaagatgaa attcagaaag ggaaggtaac ttatccaaga ttacatggct agctatgata
8100
gaaagttaga gttggaaagg acgttagaaa gtgagggttt gaaaggactt tagaagctgc
8160
ttattcaatg ttctctctgc cctttcccat cttaggcttc tccattttac ttttatccat
8220
caataaaatg ttaacttcaa aaagaatatg gcaattcttg ggtaaaagat gctctggaag
8280
tgtgagtccg ggagtattat gtgactaatg tcttaactaa gaataataat atattatgga
8340
ctagttttaa tctcttgttt caccttgaac tgttcaggaa ggaaaatagc ccacggaaat
8400
tttttaaaaa gtctttctct atctgaattg agaaaaggtg acaggcatag ttggaacatc
8460
ttttaggcag tgctggtgaa cttcaggcta ggccttgttc catgaaataa taaaaatttt
8520
caaaataatg cagaccattc ccttccaggg atgctttctc tgtaatgttt taaccccaag
8580
aaatctttct gtaaaaatct ataaaaatct ggagtgttcc aggatacaat ttgcacattc
8640
tccaatttaa ctaaaacaca atcgattttt tgttttcttt ttctttggct tagcaaggtt
8700
ttaagatagt ctctttctgg ccacagaggg agatgatttg cctctagaat accctttctg
8760
tgcttgagag agtcacaaga ctgcaagctc atggaggatg agagtcaagt agaggtggtg
8820
acatctctcc cttggccaac atccctctct ttctctttcc ttctgccttc agtggcagta
8880
gcaaaagtcc tccttctctt taggtagaca gtcagccact acaactgtgg cttcctgaaa
8940
tcctcagtgg agctatgtac ttggcacaga tttgtcttga agaagggact ccatttctga
9000
gccagttgtt gaatggggat acttagcagt acagtgaggc atttccagta ggattgttca
9060
accacaattg cccactttcc aggcccaaag gaataattga aggctatgta gacttttttt
9120
ttttttttt ttttttttt tttgagatgg agtctcgctc tgtcgcccag gctggagtgc
9180
agtggcacat ctcggctcac tgcaagctct gcctcccggg ttcacgccat tctcctgcct
9240
cagcctcccg agtagctagg cctaatatat atatattata catatatatt tatatttata
9300
tatatatata ccaccacgtc cggctaatat atatttatac ttttttttt tagtaggaaa
9360
ggggtttcac catgttagcc agtatggtct cgatctcctg acctcgtgat ccaccagcct
9420
cagcctccca aagtgctggg attacaggcg tgagccaccg tgcccgacca tgctatgtaa
9480

actttttagc agaagcttta gctattgtgt cccgaagggc cccaggtcat gatgaaatgt
9540
cttttttttt ttttgtctct tttcttctta attactgaga ctgtcaaaga atatgtcaaa
9600
gcatgacata ttccaactcc aggatccata aaacacccca agttctgtgg agaccctatc
9660
acatctgcaa aactctccag gaagtccaga gccctcctgg ttaatttgtt ttagggacta
9720
ggcatgcggt atcccctgac aacactggat cagcaattct cctacctaag tcagtcccac
9780
accatgtgca gcagagtatc cagtgcccct gccctggtct gctcacattg gtttgctctc
9840
cagaataata attcctcaat atccacaaga gattgattcc agaactactc cgaggatacc
9900
aaaaatcctc agatgctcaa gtacctggta taaaatggca cagtatttgg catatgacct
9960
aggcatattc tctcccatat actttattta tttatttatt tcgggacaga atctcattct
10020
gtcgcccagg ctgtcactcg cttattgcaa cctctgcctc ccaggttcaa gcaattctcc
10080
tgcctcagcc tcctaagtag ctgggactac agacgcatgt caccacgcct ggctactttt
10140
tgtattttta gtagagacag agtttcacca tgttggccag gctggtctca aacacctgac
10200
ctcaagtgat ccgcccacct tggcctccca aaaagctggg attacaggcg tgagctacca
10260
cgtccagccc cccatatact ttaaatcatc tctagattac ttataatacc taatacaatg
10320
taaatgttat atagttgttt taatgtattg cttttttttat ttgtattgtt ttttattgct
10380
gtattatcct tttttatgtt ttattttttc aaatattttc tacccgtggc acccacagtt
10440
ggttggtgga acctgcggtt ggtggagccc atggatgtga agggctgata gtatgagaaa
10500
actcagaggt gcagagttgg agagcacatc ggggagaatg tcagcatggg ttaaaaaaga
10560
cacactgtgg ttggagatga tcacatgaat ggccacttca aaaatgaatg ggtctcatcc
10620
tcaaagcagg ctctcctggg cactgcttgg gaaggtgcta attggagctt caggcaacaa
10680
taataagggg atacaggtgg ggatcctgcc atgggcgtag cttactttct ctggactctt
10740
ctgggtctta aggccagttt cctcatccac tcaaaagaat gacagcaagg tgagcaaagc
10800
aaggcaggta aatgaggagg actctttctg gctgtccaac ttttcatcaa cttccccaaag
10860
gtttttggat gggacatgag cactcattcc ttctccaccc tttagctagg ccctgtcaac
10920
tccaggagga aggtagaaga ggtcagagct gtggtctttc acttattcaa gatgtttcct
10980
tagtgttttg tgtttgggtt ttttttgttt tttttttttt gacagagtct tgctctgttg
11040
cccaggctgg agtgaagtga agtggcataa tctgagctca ctgcaacctc tgctttcgag
11100
ttcaagcgat tctcatgcct cagcctcctg catagctggg actacaggca tatgctacca
11160
tgcctggcta atttttgtat ttttagtaga gacggggttt tgccatgttg gccaggctga
11220
tctcaaactc ctgacttcag gtgatccagc caccttggcc tcccaaagtg ctgggattac
11280

aggtatgaac cactgcacct ggccccttat tgttggtttt taaaagagaa actaagctgt
11340
gcttccagaa cccagtttga gaaagtttga agacctggca tagagccagt gacatataat
11400
tgttagttga agaaagagag ctccttgatc tgcaaataga gcacggcccc atatttaaat
11460
tctgcacatt ctagaagcat tttgcaagaa tcaaatgctt tgaggatttt gctaaataac
11520
catggaggaa agcactagac aaatattttc agatggcatg agagttatca ttcataggaa
11580
ttatatttcc actcctacca cttactgggg acccaagtaa gaaattactt ggataagcag
11640
aggagaattt aaagttgaat gtggtggaac ttattatgga aaaaatatgt ttttctgaaa
11700
actggatatg tgtatatata taagttcagt tgtcattttg gaaccatcct tactcttcct
11760
agctaaggat tagcatacat aggtgcaact tgactaactc tgcctggacc caattcagtt
11820
acctttggt gggtagggtt catgaagaag cagttatttg tggagtgtat agaaaccact
11880
ctattgtagg ttctttagtt ggtactttca aaataagtga catccaaata gtaacttaat
11940
attccaaata tggctgcaaa acaaattgtc gattatggat gactactact gccatctctc
12000
cataccagtc catcttctgc caggctgttt ggtcttgatt tgtcgacctt ttaggtttct
12060
ccccatgtat tccacatgac cttcaccaac cccacttcta tctccaaacg tctttctgag
12120
ttgtggggat gcagatgtat tctgccacca tcacaagggc taaccgagcc ctggctgcgg
12180
atcttcattg ttgttcacat tatttccatt cttacaccct acttcatgtt tgtacactat
12240
tttcttacat ttgctgtctc ttctaaacat tctttgctgc atccactttt tctctatttg
12300
tgctctaggt gctgcagagg ctaatgctgg gtttcctttc attcctcctt gcactcagca
12360
cctcccttct caattccttt tgccatgtct ccactttaaa tcttaaccta ctccagatag
12420
tcttttcctt cacactattg gcatctgtgc ttgggttgct ttcagtctat tctctgatct
12480
atgatttctt tgcatgatca agaaggtgcc atgaaaggat cccttaagaa agcctgtcat
12540
ttagccagaa cgaactagct tcatgatagc accaggaaga ctgatatctc ccaggaaaca
12600
aaccactcat ggtggtgctc tttttgcctt cactatgaag tgtttgtctg cctgtatgtg
12660
aaaacgagag ggtttaattg taaggatgca gcacagattg ggactggcat cagaaagcca
12720
ttggggactg aggtagctct agagaccgct ttctgtctcc agtgctctcc ctcctgggtg
12780
acatgttttc tgtctcctgg catctctgct tctctctatg ggcttcttta ttatttgcag
12840
cttgcaatgg taccccaaag tcctagctca tggctcctct ctgcatatat gctttctgtt
12900
cctacccaca aagctctttc tattcttcta gtttaaattt tcaagagaag aaatctgatt
12960
tttttttaac ctggtcatgt caaagaccac tgaccacata tgagctggtt gccctgtgtc
13020
aagtgccccc ttctcccacc ctcttcccct ccccatctgg tctgtcataa ctgaatgatg
13080

gagtgggaaa ttgaaattgc catgggaatt ccatgataag ctatctaaac agttttatct
13140
ataagtggta gacagagtca cttagaaggg agtcccaggt gagacaggca cctgtcaact
13200
ccaaactggc acacattcta aggtctgcaa caccccagag agagcactga ttttgtagtg
13260
gcctgtactg ggcggtagg ctggagaatg ggagaaatag ccacttcaga atcccccagc
13320
ccaaatgcat caagctcact atagactctg cagccacgat tcagctggct tctgctcaga
13380
tcaacagaaa acattcttag tgaatgatgc ttgtggcaca tatctcaagg ctaccagggt
13440
catttcttcc catttacttt ttctctgatc tatcctctcc aggacactag cgtcagaaga
13500
taatcttccg tcgttttcag gtacactatt tgggtactga gtcactttca aagcctcttt
13560
ctgggtttgg atttccagag cagcctgtgc tgtaaagcaa gacagaaagc ttccctgcca
13620
ttcatgcctg ccagggatag aatgacagta ctcctgaggc tctccctccc cacccctccc
13680
ctgctggaca gctgatctgc tggactcagc cagagccagc aggcaccccc tctttatcct
13740
aggagctgca aacttgatgc ctttccagga aatccccaga agctggagta tcctcatcta
13800
catgtggcac agtgtatggt tgtgtcaggt gctcatgtcc cattgcatag gactggggtg
13860
gaaaataggg accgtccttt tgtgtcagct ccagtcaatg agtagtggcc atccaggggg
13920
ccatcttgga aaggacttgt gaggctgtat ctgcgctcag ttgtagatgt gagaagaaaa
13980
ggccaaatat ctgccaatcc tagtcctggg attcaagata gaaagaactg catggagtga
14040
agaaactagg agtctccatt tcactgagat gcataagaat gaaattattg tcactatttc
14100
ttcaatactg ggccaatcct aataagaaaa ccctttttga gtctctcttt tctttatcct
14160
acatataaca cagaagcttt ttctattccc tggatgaacc cacagggaca gaaattcttg
14220
ttggacaggt gaagcagata atttctttat cagactagaa tcttccagaa gcactgctaa
14280
cctagtgagt tttgtactct agacaggtgg ttctcaagcc agctccccac cgcaggcctt
14340
tttcatggtc tgcccctccc tgtggaaccc atgttttagg ttattagctg ataattggat
14400
ttctattttt tctcataaaa tacagcaaaa gatagctagt gatattatga tgagttaatg
14460
taattatagc caaagcagag agaaacaaca ttttaattaa cctgtgtgga ctgctggaag
14520
aatataaact ttctattttg ggggttgagt agagacagaa atgaacacag ccaagggctg
14580
actgtcagag gacatttaac tgatgtaaaa tgctttgaaa ttattgggca ctcattgttt
14640
aaagttgttt ttgatgatgg taactccgta aggggatcag aacatgctgg aaagaatggg
14700
cacagctttg gttacctggg ccttaccact gttattcagg cctctgagaa agcttactat
14760
tgttgttatg tttcttacat aataaaactt ctaatatttg tatgaaaaca tagaattcca
14820
cttttaaaga tgtaaggatt ttgtcatacc attagggtta ctatgatcac ttgattctag
14880

gtctaagaaa tattaagtaa tttacccgcg aacacagagt tttaagggta agtatcaaaa
14940
ccttgatctt ctaataccac atattctcac tcatatgtgg gagctaaaaa tattgagctc
15000
aaaaaggtag agagtagaat tgtagttatt agaggatgcg aaggaggata gggagaggtt
15060
ggttaatgga tacaatgtga agttatgtaa gaggagtaag ttctagtgtt ttgtagcact
15120
gtagggtgaa tatggttaac agtaatttag tgtatattta aaaaaaaata gacaggattc
15180
tgaatattca caaagaaatg ataaatattc agctgggcgt ggttgctcac gcctatattc
15240
ccagcatttt gggaggccga ggtggatgga tcacctaagg tcaggagttt gagatcagcc
15300
tggacaacat ggtgaaaccc cgtctctact ataaatacaa aaaattagct gggcatggtg
15360
gcgcacacct gtagtcctag ctacttaaga ggctgaggca ggagaatcgc ttgaacctgg
15420
gaggcagagg ttgcagtgag ccgagatcac gccactgcac tccagcctgg gtgacagagt
15480
gacactctgt ctcaaaaaaa aaaaaaaaaa gaatgataaa tatttaaggt gatagatatg
15540
ctaattaccc tgatttgatc attacacttt gtatacatgt gtcaaaatat cactctgtat
15600
ccatacatat gtataattat tatgtgtcaa ctaaaaataa aaggaaaaaa atcatttcag
15660
tgtatttaca aaacatatgt aaccattaag aataatgttt taaattatat ctaagggtgt
15720
gataaaatta cagtataaga ttgtgcttga aaaagtgcaa taagaagtaa atatgtacag
15780
atgagaaaaa gtgcaaagaa ctaagtccta agcagactat acctttccta ctgcatggta
15840
cttctctggc cttttgcttt gaaagatttt gcacccagca tggcaagtgg ttagcagagg
15900
cagccattct cacttgtgcg ttggctttgg gagccatata tgttgttcag ctgggtgtgg
15960
agtggaaagg ctgcatgttg tattaatgca ttgttaagaa cctctaagag tgatttcttt
16020
tgggaagtga gactgacggt ccgaatggtg gaaagacaac ttttaatctt ttactttaca
16080
ctttgtgcac ttttaaatgt ttaacatgag catgcatttc tttaataata aaaatacaaa
16140
aaaattttag ccctagatct tctgatttta aactgcatat tctttctatt gtgttacata
16200
ttttagcatg agaataaggt tatgaagctg gaagtagcag gctccctttt cctcatatgt
16260
aggaagttaa gaatgcattc tacgtttctt ctttaaggag ttggcttctt tccttttaac
16320
atagggggtaa ctgggcccag ggagtttggc aagggccaaa taaagtcctt aatgcccagc
16380
tcagaaatct ggattcacca tccttgactg ctggctccaa cccaccctca cctgagctgg
16440
tctgcagagg attcttgttt gtgtcacttc atcaccagca actaccgaca gatgatgctt
16500
tggcctgctg cctgggtaac agggcgaggc tggctcagga ccatgttttc agatcagggg
16560
acctcctttg atgccatgtc catggtgtcc gagggcagcc aggatcaagg gctagacggg
16620
gcagtgatga gatgagagca ggaggggctc agctgcagcc ccaggagagc ctatgccagc
16680

cctgttgacc aaggaggaca gaagcaacag gagagcggag gcagaggggt gagtgtctat 16740
cgctcaatgt ataatcggca gacatttggg gagctcatac tgtgggctaa gcacagggaa 16800
gaaaggcaca gtccctgtcc tcagggaggt cacagttgat agggaagaca agcatatgtg 16860
ctagctgcta tagaaggggg aaccactgag ggctgtggcc acacagaggc aacaccccct 16920
tcttgttttt ttgtcaggga ttcagtttgg cgtcattaga agtgacttgc acaacccccct 16980
cctccagtca attcagaagg acttgttaag caggaatgat gaattagctt cagcttgtgg 17040
ggcacacaca gatggaagta taaggtggcc tcaggagtaa gtaaatcccc atgcaagctg 17100
tgtccttaga ccagagcagc acccggttct tccccatttc tagtaaaggt gcctcacaca 17160
ccaccaggac acaatttatg cctgcagaat gaatgaatga atgaatgagt gaattcctgg 17220
aacctcttct gcttatgtgc cacaccaggt tgcagcaagc ccagggacac ctgggactgg 17280
aattgggctc tcaggtgtaa ggaccaggga gcacccacca ttttgcattc ttcagccctt 17340
cctcctctcc tgtcccagct tcagcaatat ccacagagcc ctctgagcaa ctctgagcct 17400
ctccacagcc tgacgcctgc ctgggcacca gctcttcaga gggtgtttct gtgctgctca 17460
gctacctctg agcctgggct gcctttgatg ctcaggagac accctgtaat tcaattaagc 17520
cttctctcca gggagcatgt aattatgtcc tatctgggcc ttgtaatgac agccccctgc 17580
cactctacag ggagttgccc tgctcagctg cccagaacct ttccctggga ggaaactaat 17640
ctgcttagcc cagattggac gcagttctgc acagcacttt tccgaatgcc tctgaaatga 17700
gtcctcactg acagaacggg cccactctgg gggaactgag ggctctcttg gtcctgcact 17760
gctctttgcc atacagatct gtctgcccag gatttttctt gggtgtgtag gaggctgaga 17820
gagctcccct ttcttctcat ggctaaatcc cttggtcttt ccagccctcc tgggggttag 17880
aagggagagg gaaaaaaaaa aagactgaac ttgttgttgt tgtttttgtt gttgttgttg 17940
tttgcctgtt ttctatgttg tcttgtgggg agagggtata agattgattg acagagtggc 18000
acacttcccc tgcaaattca tcatttgaat ttctcaggta agatgttcac atttctctgt 18060
taagatgctc caatttctct ggttaagatt tctctggtaa gatgctcatg aattggtgga 18120
ggtgttggcg ggatgtggga agtgtgcctg ctctttctga gttttggggg aagttgcctt 18180
aattctctgc atgactttct ttgctccttt gggcttcatt tctgtgcaat gtagtctgac 18240
atgaatactg ctcagggagg tgttgcttcc cactgcccac gccactggaa accagtagcc 18300
caggtttact cgagtcctcc ttttgaggaa cccaaattct ttcatttctt ttatgtgaga 18360
tctgcccaaa atgccattgg caagctgtac tgggttgaat agtgtccttc ctcctcccaa 18420
atgtatgtct actccaaacc acaggatact accttatttg ggaatagggc ttttgcaggt 18480

gtaaccatta atagttatga tgaggttata ctagattaga atgggcccta gatcctatga
18540
ctggtatcct tacaagaagg ccatgtgatg acaaagacaa agaatggagt gaggcaccca
18600
aggaactcca aggattgcta ggaaacacca gaagcttgga ggaaggcatg gaacagattc
18660
tcctctcgga cctctagaag gaatcagtcc tgctgatacc ttgattttgg acttctagcc
18720
tccagacctg ttggggagaa tacatttcta ctgttttaag ctaccacgtt tgtggcgatt
18780
tgtcacagca gccataggaa actaatacat acaacctgca caatgcctac tccagcattc
18840
catagcaagt caagggcctc acaattatgt ccaaaggact gatagaagag cgacctctgt
18900
gctacttgtc cctcaggacg ctgacccaca gctctcaagg caggagtagg ccagagctca
18960
ttcaacaact ttgttatata ggggttccaa ttgtaaacct tttgaattcc tgtttgcaag
19020
tagatgaggg ttgaaaaata aatggccact ttctctaagc cacatacccc aatctgtttt
19080
gttacttcat tacagctgtt ataatggcct cctcttctat cttccaatct ccatagccct
19140
ggttccttga tagttctttt tttttttttt tcttttttttg aggcggagtc tcgcactgtc
19200
gcctgggctg gagtgcagtg gcacgatctc ggctcactgc cacctctgcc tcccaggttc
19260
aagcaagtct cctgcctcag ccacctgagt agctgggatt acaggcacct gccaccatgc
19320
ctggccaatt ttttgtactt ttagcagagg tggggtttca ccatgttggc caggctggtc
19380
ttgaactcct gacctcgtga tccacccacc tcagcctctc aaagtgcggg gattacaggc
19440
atgagctacc gcgcctggcc agatagttct taaacaactg cccagaagtt ccagcctagg
19500
caggggcagc catgaactgc attgctcatt tctgcttttt gaccttttcg atggctgaac
19560
tctaggccat ggaaaacaag gacccactgt atagttaaga gtcattttgt gactagggag
19620
acaaaaaagg gcctattctc caaatcccct ttccctctgg agttcctcgg tgccttaaag
19680
cttgtcctga gctacaggtg tgttacctgc ttatcccaaa atgcaggcat gttacctgct
19740
ttcctctgca aagagaggca ggcctggctg gggcacagct gaagatgtca aggccaacct
19800
aagggcagcc aagctatggc tgtctgtgac aagaggagag cagcggtgat gggagggtag
19860
gaggcattga gttcatgtcc gggtttgcct cctaccctcc tatcactgct tgatgatcct
19920
atcactgtct tgatgagttc aagacagaag tttgcctcat cattgccaca ataaaatcac
19980
caataacaga agtgtgaaag cagcgatgtg agtggaagcc catatataca caggggggtaa
20040
tagagcagca tgattaaata tgtggccttg ttatcagaca ggctgatttg gagtcccagc
20100
tacttgttgg tgacctgaac tagaggaagt tatctaacct ttcattttac tcatttacat
20160
aacatggcta ataatagcac ctaccttata gggttattgt gaggattgaa tacaattatg
20220
caatataaaa cgtttagcat agtgcctagt ctaaattcct caccaggggt atgatgtact
20280

agtttttagt taagtaatta gtatcctgga catgtcacag ccatttgacc tatctgggcc
20340
agcgttttgc tcaggttccc ccagcagtaa ttgtattccc tccccaatcc cgggattagc
20400
ttttaggaag aaacagttga tctaaagata gaaagtcaga gtactgtctg gaggaaggta
20460
gagggaaatg tcattatctg ggttttcttt gatgatgtca gggaacatga caggctgctc
20520
ccaaagacag agcagcccca ggacagggaa gaaggtgacc ttgaggttga ctcctctgca
20580
tcccgatgtg gacgttatgg acttgttttg gagatgaagg gaaagaaaga tggaatgtag
20640
aaagtgaagg agaataaaag aagtgggagg aagaagggct gggaggagga tgggcaaagt
20700
ctttctggtc tcaaggataa ttacatgtga aatcacttgc cagtgggact ctggggctgg
20760
agcagctaca ataattacag tacaggctgc agagggctct tgggcatgtc ttggagcagc
20820
ctgtaggcag tactgaggcc tctctcacta gacccatctc ccagatcaca tagtacacac
20880
accttccacc cccgggcctg ttaatgatca aaaagcttaa acagaacaat tacagcttca
20940
gagtggaacc atatctctgg gctcctgtga tgaaaaccac aagcctgtca ggctggggct
21000
gcttcacatg gagggccctg ctcttaatgg ccaagtgatc tggagcaaga cccgtgactc
21060
tcccatagtg ctgtggatgg tgctgcctct ccccacgcat ccccagaaga ggaagttcag
21120
taactaagga attaactatt ctccagcctg attctgcttt tcccaatcag ggctttatac
21180
ctttcttttt catccctata tttggagatg agtcaccctt gccttcattt tacctaagca
21240
aggcagtttc ctgtaaccta atgaagtgcc aaacaatact gtgatttatt tagtacttac
21300
tgtgtgccag gaattccagc aggtgttgga catttatgat gtatgatcct tacactaagc
21360
ctgcaatggt gcaaccccag ccctgaccac tctgtgcttc ccttttcaca acacagcttg
21420
tcactaaatc caagtcagga attccaggtt aggcttgagt tgtgcagagc ccttaactga
21480
aatttgccat ggttgaggca tgattgcaat cactgacaac tcctcccggc tctacacacc
21540
tacttgtcat attcacgccc tgatcacggc cccactcgca tctcttccca ctttagaagt
21600
tctttcctat agaacacgtt gctgctgccc tgttctggtc actgatcagc cctggcctaa
21660
ccactggcta agctttgtgc ttgcacatag ctggttgaat cgtatgtatt gctgtttgtg
21720
tacatcaaaa atataataat aatatcggca attttatgtg tttcattcaa catgagggac
21780
ccagcattct taccttgtcg ctttgtaaac cctgctgctc tcaaatctcc actagctgtt
21840
tcctgagcag aaggagataa aaggctggct cacacccccca tgtttttact ggtcacagtt
21900
actgccacca tccaaggctg aagagacttc ctttgtgtta gggctaaaac cttagtcatt
21960
gtatctaaat gtcttctgta ttcctttcct caaaagaaaa aagtaccctc ttctgccaac
22020
cctctcccat gccaactaaa caagcaagca agcaaacaac aaagaaaagg tgatattaca
22080

gatgctgctc agcctatgat ggggttacat cctgataaac ccatcacaag ggatgtaatt
22140
ccattgcaag ttacaaatac cataagtcaa aaatgtattt atttcatata acccacagaa
22200
cgtgatagct tagcttagcc tacttgatca tgttcagaag acttatattc gtctacaagt
22260
ggacaaaaac atataaaaca aagcctattt taaaataagg tgttgaatat ctcatataat
22320
ttattgaata ttgtactgaa agtgaaaaat agaatggttt tctggatact caaagtatag
22380
tttctactga atgcatatca cttttgcacc atcataaact tcaaaaattg tcggtcgaac
22440
cttcctgagt caggaatcct gtctgtacag ggtataaagg aggaaagcat cagctttgga
22500
ggcaggtgga cctgtgtttg aaccctgatt ctgctagagc ttgacaatgc atattcgttt
22560
tctattgcat aactaattac tacaaacaac acatttattt ctcagttttc atgaatcatg
22620
agtccaggca caatttagct gcagttaagg tgttagctgg ggctgctgtc ttatctgaag
22680
catgggggtg ggggtgtgga ttccaaggtc aggtggttgt tggcaaaatt aattttcttg
22740
cagctataga actcatggct tgcttcttca aggacacggg gagagagaat ctctcacatc
22800
ttttaaaggg ttcacctgat taggtcaggt ccactcagga cagtttccct taaagtcaag
22860
gcttaatagt caactgatta gggaccctaa ttatatctgc aaaatacctt caccattgcc
22920
atgtaacata atcatggcaa ataatcacag gtcccaaatg ttcacaggtc ccactcacac
22980
ttgagggagg ggattatata gggcatgttc ttgcggagag aaggaatctt acagccacat
23040
tggaatctgt cttccatgct atttgacctc aggcaaattg actaatctct tgaaggttca
23100
atttccttac ctggaataaa aggacaataa gatcagccat ataaggctat gacaaagact
23160
aaatgagata gaataggctg gaaaagtctt gcagatagca gacacaagta tataacaatt
23220
tccctcctac tgttcctttt gtttttcacc tatcctgcag tctctgtcac ttcaaatacc
23280
atagaaaacc tttccaagca gcccaaatca tgccccccaaa tagtcacgtc tcattattca
23340
tagcagttat gttccataaa gttagcacaa actccgaatg agtgaatcct aaagcgttgc
23400
tcctggagga aatacaggct gctggtcaca atatttttat caactgatca atatataccct
23460
tgtcttatgt gtgtttctgc ttcaagacac tttatttaat atatacgttg attcattaac
23520
tctgaactct ctaggcaaca gcattataac tcctgccttc acaaagctta tctaacacac
23580
acatttcctc ctcaggcaca tcccagcctt cttgcactta ggattcagca gtatgcttaa
23640
gggccatttt caacagcaaa ctcatcagcg caaacacaaa catgtgaaaa acgtagcact
23700
aaagagactg caaaaaggac actggcttac agcatggaag ctggaaggag aaggcagaga
23760
atcaccttgt tccacttcag ctatgaatat gcagtcaggc cacccagtca ttcaaatttt
23820
ataaatatac tctaatatat atataaatac caggcagggt tattttttc ctcaagtcat
23880

ttttctaatt ttttttaaat gaatagatag aagagctgaa gtaagggtca ggagcaagag
23940
ctctgcttcc ttttcccttg ctgggcttcg ttagagagcc atcatctcct caatatgtct
24000
cccaactctt ctaggcattg gatgagtttg ctgcagatac gaaacccaac tttgccagtc
24060
acttcatact aacaggtgaa atgtagtgga ggagcctttt gaagacaggg actcagcccc
24120
ccattagcct cattgcagac ctagattcct gccaaaatta atttggctgg aacttcccag
24180
ccatggcatt gtcgacatta cacatcttcc actgtaatgt caattaccat tttattcagc
24240
cgaatgctgg agagttaatg ttcaagtggt tagagctggc tacgggtggg ctgaacaaga
24300
tgtcttttcc ttcatttccc ctgcctgtgg tgaaggattg taaccagccc tggctggcag
24360
cactttgaag ctcacccaga gtgctcctgg ggacatcttc tacagagcct atcatttgga
24420
catgctgtct tctgggcctg tcttccttcc ttccttcttc cctccctccc tccctctttt
24480
ccttccttcc ttccttcctt ctttccttcc atctgcttta aaaccagctg ccttgagtgc
24540
ttgtcttggc gcccctcatt agtgccattg caatcatccc tcctgcctac cctgctaacc
24600
acagcttgtt agtccacaac agcaacagct gtgtgctggg gtgcagcagc tggagggcca
24660
aaggtagggc tgggggacag ggtgttggga tggttttctg gggcagatga gtttatacgt
24720
ttctttcatg tccccttcct cccacataga cttttatttc cccaaaggaa aacagaaaac
24780
aatgatctgt ttgacagtgt tgctatcatt gggcatcaaa cctatcatct aaggggaatc
24840
cccctgtata atcagtcagc caaatggagc aggaccctgt gttttgtagc tgatacaaca
24900
gggcagcatc tctagtgagg gggccagggc ttctatttcc ttcattaaaa aatgaaacag
24960
cagacctgat tccatattta gagattacac ttagttgcca ctgtgggtgt gcaggcacca
25020
accaaaccca gttggcaccg ttgtcttttc tctgcaatga tgtattgaat ttaataatgg
25080
aggtatatga aattcagagt gattggaact gaaggtttag gggctttgtg taaaattgat
25140
atgtaaggga tttggaagta ggtgagggat tcttccccaa tacttattca attttggagt
25200
caaataacca agcatttaca aatagccaaa aaagaaattg aaagagggtt taatccaata
25260
aattttcatg cctcatatga accacatctt ataataagaa ttatgctttt tcatttcata
25320
ctcagttaac aaatatgatt tgtgagcacc tggtaagttc agggcactag gctgaaaggg
25380
gttaccaaat gtcttcattt aacaaagtcc agctgagctc ttacaggtac cagaactgtg
25440
cctgggctgt catatgaaga tgaatgtaag agtgtgtcag gccttcaaga gcttacagtg
25500
tgtcaggaga catcaaacaa gtgagccaat aaaatgatac tgccatttta gaaatagcct
25560
gaaattcatg gagttcacag tcttgttagg aaagtgaaac ataaacctat aagcattaaa
25620
aaataactgt tgaagacagt aacggaagaa tgcaactggc aactgaatga tataggttgt
25680

gatgactgtt aaatatcatg aaaagagacc atgatgagct gaggcactcc aagagacttc
25740
tttttggaga tatgtttgga gccaaatctt gaagatttaa ttgcttttt ctttttttt
25800
tttttaggtg gagtctcgct ctgttgccca ggctggaagt gcagtggcat gatctctgct
25860
cattgcaacc tctgcctcca ggttcaagcg attctcctgc ctcggcctcc tgagtagctg
25920
ggattacagg cgtgtgccac catacccagc tgatttttgt atttctagta gagatggggt
25980
tttgccctgt tggccaagct ggtctcaaac tcctgacctc aagtgatcta ctcgccttgg
26040
ccttccaaag tgctgggatt acaggcatga gcactgtgcc tggccttttt ttttttttt
26100
ttaaaaaaaa aaaaaaaaaa aacaggaagt tttcgttagt ttttttgttt gttttacttc
26160
ccataaaaac tctttgtgtc acatggaggt gaatggaaag agaggctgtg gcaacagacg
26220
ggagactttt ctgatatcag aacccagtcc catagaccag aatgtatgct ttcaatccac
26280
gttgtctggg tccatcctat tgagtgccct gcccccacag cggggtatgg agaagagtca
26340
gacacagccc cagtcctcac gtagctcaca atccagtgga ggagacggac tcagaaacag
26400
atagagatga agccatgaga tcagtactgt ccgaggccat ggccacggtt ttgtgggaac
26460
ccacgagagg gaatgactaa ctgtggggaa gaagagggag aggaccaaaa tgcaggggaa
26520
gtgctcacag aggataagta agcagtgagg tgccatgaaa tgagtataca cctgacagcc
26580
gtgtaacagc tcagagcctg ggtagagggg aatagagctg ctggttctct gggggggaaga
26640
gaggggtatg ggattctgga acagaagcac caaaaccagc aggttattgg agctgttagt
26700
gctcagatca gcaatgggtg cacaaccaaa ccattctcct agggatgagt tctttcctgt
26760
ggatgagggc ttctcagcct ggcttctccc gagaattacc cgggaagctt gaaaagtact
26820
gatgcctgga acctacctcc agagagttgg atttcattgt gttgacgtgg ggctgggata
26880
tcagtatatt gtttaagcac tccaggtgat tctgatacgt agctgtgatt gagaaccctt
26940
gccctaagct atccatctgc actccagggg tgctcccagg cccatctgtt tgtaaatgga
27000
caggtgtctt gaggtaacaa atgtgccaag gctctggagc caagcacgcc tggctcctta
27060
gtgcctactt agtgacctca ggcaagttac taaatggctt aaactttaca aatccttaat
27120
ttgtaaaatg tgggcaatga tagtacctcc tcacaggatt attacgaggt ttacacggaa
27180
tactctcagc tcataataag cacttgcaca ggcctcatgg gctaggccct caaaacttaa
27240
cgcatctaca ggcaacagcc atatgaaagg aattttatac caccaagtca aaaaatctgt
27300
gagcactgct cagaagcaaa agcctgtctc caacagcgct catttaaggg gtgggcgagc
27360
tacagagaga agaatgagcc cccacagggt aagctgggga aagctgggga cagaatgaga
27420
ctcaggaaat cacttgaata ttgattatat ttgtgctcaa taataaaata acgaaatgag
27480

tacagcccta gacctaaaca ttgtgggtga ggcaaaggca atgcgttaat tttgcatcca
27540
ctgaggaaaa actctaaaac ggtgacttct tttttaaggg accagaagaa tctagattat
27600
atttagtcta agtcaataca tacgacagaa ccttgccctc tagacttgat aagaaagaag
27660
taaaataaga gaaagaataa aaaacccttc caccaaaata ctaacattca gataatgact
27720
ttttagttag gtctcctgga gaggaggttc cctcagaaat gaatagattt ctcttctagt
27780
gcaatcatca aaaggtaatg catggactta agtgtgatcc ccaagagaaa atcaatgacc
27840
tttctgtgtt tgcctttgag aaaatcagcc agtctatggt taaattagac atattttttc
27900
tccttggtca agattagtgg gaccaagaat gcagtcttac actccttcta gcaaagaatt
27960
acctgatgcc ttatttcaca caaatttgca aagttgtatg gacgttgtat cttattttaa
28020
ggagaactgg tgatcaaatg atgactattt caatagtggt tcatttacac caccaccctc
28080
accccacatc ctgctttcac ctgaatctga acgatcatag tcagtctgag attctgaagg
28140
tttgaaattc cttttctgag ctctgcaaga acagcatctc ccaagagagc tcagggcaga
28200
ctgtctggga gagattggaa acctgtcttt tgcagtaaca tgaattggtt gaatggtcac
28260
cctccatatc aggcctgctt ctcccattgg gtttctgatc agcccaactt gggtctcacc
28320
cttctgattt ctctctcctg gctcacatgg ggctgcactg gccattaggt gccaggcttg
28380
gctccgtgga acccattggc cagctgggct ctgtggagcc ctaaggcagg gctctggtca
28440
ctggtgagag ggaggccatt ggagtcactg gggtggacct acagacccta gggttaacag
28500
ctaggtgggt gtcctcttca gagaaacggg ttacaaagtg aaagaaagtt acactgtgag
28560
gtcagccagg gaggaagaca gagagctgat ataagatagg tactgattcc ctggggatgt
28620
gaaggaggg taatattcct aaaatgatag catttagctt ccagtataca ttaattgatt
28680
cctgatattc attaaaacta aacgctattt ccttgatgtc tcatccaaag ccgcaccact
28740
cttcccacta agtctgaggg gagcttgttt tgttgacaag tgtaagaggt tgaagaggga
28800
cccatgaact cttttgtcct actgaagaga tccacagatg gaaacaaatg ctcctaccac
28860
atttatgaac tgctgctttg cagtcccgct tctgctatca tgcacaggaa ctgactaagc
28920
tccaaagcca gaggatgtaa atctccctgt aataaatgta agtcatttat tagctacata
28980
cacttcagca agtcacctaa cctgcaaatt tcaagcatgt gaatcttgga tctttcatgt
29040
gctagctgtg agactttgag aaatgtattt aatgtctctt tgcttccttt tctacccaca
29100
caatgggtat aataatgtct accatatatc tttgcagcaa ggtctaaatg gggtgataca
29160
tgctgaatac atttccaaca gagtctgtgc aatgataagc tctttccaaa tgttagttaa
29220
agctaaccaa ctaacccacc aacaaaccaa cctcttagcc aggactgatg gaaggagtct
29280

gtgagagaat gcatttaaaa cacttggcac catgcctgac aagagtaagt actcgataaa 29340
tcagttattg ttattatcgc atcggtatta tgaccattat cctcttctct ataggcttca 29400
ggtttcctg tctttttatc acagcagtat tccagcagaa gcctttgatt taactaagtc 29460
tctactgtgt gtgtggctag atgctataaa gcatccagag aagtgagaat ttggtcctgc 29520
ttttaagtag cttatagtct aattaggggg aagtaatcag atagaaagga aactaacaat 29580
atgcaaaagg aaactcatag tttgtggtaa atgccaggtg ctgctgatag tggcttcaga 29640
gagatctcat agatgctata ggaggtcaaa ggagaagcgt gcagcttgag ctaagttttc 29700
agggaaaagg gtgaaagaat tagtcattaa tgtacaccta cattacctgc cagactccat 29760
tcaaaaatat tcttaccaaa tcatcacaat accttgttgg taggtactat tactatttta 29820
cagaggagga aagtgaggca aagacacatt aaataatttt cccagaatcc caaggtgtga 29880
ggtggagcaa ggacacaaat ccatggctct aagtccctcc tagtatatcc tgcaaacaca 29940
tctggaatta atgcagagag gaaggggaga ggcagtgttc tgcaggagtt cagagccatg 30000
ataacccttc ttgtgtggct tttggtaagt tattttacct cttaccctct gtttccccat 30060
ctgttcaatg aaggttgtat atacacacat tatatggccg ctgtaagtgt gcagtgatat 30120
gatgcatggg gactcagttc atgaggcagt gtgaattctg aaggtatcac aatgggacag 30180
gtgttttttt ctccactcat tttctccgaa agtcttttgt tttgttgccc tccctctttg 30240
gggcatatgc tttcagctca taccttaatg acatcagaat ctgcaatttc ctggcaactt 30300
ttgtggttaa aattattctg cccttccatt ttaaagcact aatagcaaag gtattaggtg 30360
caaaatgatg ataaaaataa ttgcaatttt taccattaaa agtcatggca aaaccacaat 30420
tactttggca ccagctgaat attttgaaac tccctactct gatgttaacc aagttcatga 30480
ttcaaagaac ttgcagaggg gtaggggaat ttcaagggaa agggggagat gcctggggtt 30540
gtcacacact ctgtctttca tcctctattg acatgttggt tatttggaga tggtattcag 30600
ttccactata gcccctcagt cactgtagac cctctcaaag gggcaatcat gtttccctta 30660
ggtcaggtcc attcatctaa cccctctccc gggggcatca ccttgtttgt tccagcagct 30720
gtctggccaa actcacacct cctcctcacc ctctagccct tatgatctgc tttggggagc 30780
catgggaacc cctagtttcc tctttcatac ccactgagat tcacaagtaa ctaaggtcaa 30840
ggcggggctt cattgccttt ctgcagatac cttacgctac tgttcctcct cgcctggctg 30900
gctccacact ccagcagacc ttctgctggg cgagaagctg caggcctgaa tctctgtgtt 30960
ctcatatggc cccaactctt gggattacac tagctcttgt aagaactcaa tgctctgctc 31020
tgctcatttt gatgccatca aagagggctt gcaagttacc agctgggagt gaacaccagt 31080

gtcctctttt tagaggtacc cctaatcttt ctgaacaatt ttgctggcac cccttcactt
31140
ggctttgccg gggtaagagg gggcacttct ctcctttccc tcatgaaagg agggagagaa
31200
gccaaaaatc tccctactag tcaacaactc aggcacccct ccttctctcc tctattttat
31260
agactgggaa gggagtgatg gttgttggag gtggcagagc cagttcagct gccttttgtg
31320
aagtcctgaa ggaggtgtct atcctcaact gctggcttct gtccttaagc ctggggagaa
31380
ttaagtcctc tttgcctcag tttggcactc caattgccaa cattgggaca gcaggaaaag
31440
ttccatccaa catcccatta aatatgtaat gtgtattagc acagcgcctg gcactgggca
31500
ggtattttct aagtgatagc caatgcgaag cctactttat tattttcctc tttgcttaac
31560
ctacaaggtg tctaagacca tttgtttgtc cacacatagt aagataaaca gcactgagac
31620
tgtggtcctt tctgccctgt gtccttatcc cacctgggaa tctggaaagc caagcctaga
31680
cacactcgtt ccacaaatgt ttactgaagc ttgttctatt caaagcactg tacagctaca
31740
aagaccatct tttctgaact ccaaaccagg ccacatggtt ggaataactt caagtatgga
31800
gaccaagaga aaaggtggtt gttgtcagca aagctctgag tccacacctt ccaggaactt
31860
atagttgatg caatggtggg agaagtctga acctggattc aatctgcttg attccgatga
31920
atggtgcagt aggcagagcc atgagttcag agcaggaaga aaccactggt tcaaagaagc
31980
atctgtcaca tcgaagctgc tttatagtct gttgggaagc atgcataata atttattctt
32040
tctttctttc ctttggtcaa caaagatttc ttgagtccct actatgtgcc aggtactctt
32100
ctaggtactg aagatgcagc agtgaacaaa gaagatacaa tccctgccca gcggagctta
32160
cattctagtt atcgaaagtc cctttctcag tggctgctct ctttatttga gaaaccatgg
32220
gctgttctcc tcccatccta gggctgctgg ctccacagag gcacacagtc catcaggatg
32280
ctctgccagc cacccaccca ctcaagacca agggttacgc tgtcagtgtg agcagggaca
32340
ctccccgtctc tgctacctcc tttctcctga aaacaagatc tcagggaaca tctgccatcc
32400
atttccctc cctggggagt gacaggaaag gtgtatggag gagattgagc ggagtgatgg
32460
attgaggcac tgtgaaagtg aatcattgcc tgacatggga atgaggagac ttgcttaaag
32520
gacaagccat gctaagtcat ccatcgttct cccctaagga ggtgaattga agttcccatt
32580
tttcccaggg agccaaatta acaaggtgct gggagatttc caaattagaa aaaaaaaaaa
32640
aaaaaggcac caccagctct caaatcagag aggctgttga gttgtttttt ggagcagatc
32700
attgtatttg gcatctaacc ttgaaataga ggagaaagca tggaatttct gctgaaaact
32760
catccttctc tgagcaggtg gtacaaataa gcatcgttgt gttctcagag gcaggaacca
32820
catttgcacc ttgataccaa ctacctcaat aaccacagtg ctgaatttc acaaattgcg
32880

aattaggaaa ttgttgctca ttttacaatt tggtttccct caggattcct tttaagtagc
32940
cagctacccc agtacttttg aaatatgact tgcttataaa aatttgatag gcttggcacg
33000
gtggctcaca cctgtaatcc cagcactttg ggaggccgat gtggggtgga tcacgaggtc
33060
aggagttcaa gaccaacatg gtgaaacccct gtccctacta aaaatacaaa aactagccag
33120
gcatggtggc acatgcctgt aattccagct gctcgggagg ccaggcagct aggcaggaga
33180
atcacttgaa cccaggagat ggaggttgca gtgagccaag atcatgccac tgcactccat
33240
cctgggtgac agagcaagac ttcatctcaa aaaaaaaaaa aagatatata aacaagtttt
33300
tataatattc tcaatatgaa ctagtagaaa aaaagcatgt gtttttaggt cttagaggcc
33360
tggttcccag ttttatctct gactctaatg aggtatagta ttacctacat tgattagccc
33420
ttctatactt cataggagat gctccaagac tgctagcttt cttcattcaa taaagagaga
33480
tataacagga tgggccttaa aagtagcatg catttcttct ttcattcact cattcaaaat
33540
attttcatgc gtgaaaatgc caaggatgtt tggtcaacca actcttccca gaccctggct
33600
gtgagcctgg cttagaacaa ttccatttta atggtccatg ccctcaggca cttgtattct
33660
agtagaagag caaggtaaga aaacagctta aaaagttaaa cagttttagg ttgagatggg
33720
tgttgtgaga aaaataagca ggatgctttg aacctatgca ggtaggaagg tctggaaagg
33780
cctctctgat atggtgatgg ttaaagcaaa accaaaaaga ccaagaacac atggaacaca
33840
tgaagggctg gaagaacagt gttttatggg gaaggactag tacacacaaa ggctgcaaag
33900
gcgagtgggc tcattatgtt ctagaacatg ccaaaaagcg ggtgcagctg gagagggagt
33960
aagatggcac aaaaggtgag tgaggtggac aggagcctta tcacgcaggc ttacacaggc
34020
tctcagaagc cctgcgtgtt ggtttcttgg gactaccgta acaaagctcc acatactggg
34080
tggcgtaaaa caacaaaaat gtattgcctc acagttctgg aggccagaat tccaaaatcg
34140
ggtgctggca gggctgcgct ccctccaaaa cctgtagagg agaatccttc cttgcctgtc
34200
cctagcttcc agtgggttgc tagcaatcct gggctgggtg actccagctc tgccttggtt
34260
gtcacagggc gttgtctttg tgtgtctctg acttcacata gccctcttct tcttcttttt
34320
gtgtgtgtct gtgtgtgtcc actctgaggc acagaagttt ttatttattt atttattcat
34380
ttatttattt cattgataaa cataatagtt atgcatagtt ttggggtaca tgagatattg
34440
gatacatgtg tacagtgtgt gataatcaaa tcagggtgat tggaatatcc attcacctcc
34500
aaacattttc tcatttcttt gattggggac attataattc ttctagctat tttgaaatat
34560
acaatagatt attgtttact ataatttccc tgctgtacta tcgaatacta gaacttattc
34620
cttctgttga gggtgtactt ttgcacccat taaccaactt ttctttatgt cctccttccc
34680

acttccctta ccagcctctg gtaaccacca atctactctc taccaccatg aaatcaactt
34740
ttttttttta tagctctcat atatgagtga gactatgcag tgtttgtctt ctgtgcctgg
34800
cttatttcac tcaacataat gacctccagt tctgtccatg ctgctgcaaa tgacaggatc
34860
ttatttattt ttttatggct aaatggtatt ccattttgta tgtatatcat atcttcttta
34920
tccattcatc cactgatgca tatttaggtt gattccatat cttggctatt gtgaatagtg
34980
ctccaataac catggaagtg aaaatatctc ttcaacatac tgatttcctt tcttttggat
35040
atatacccag tggtaggatt gctagatcat atggcagttc taactttaga ttttaaagga
35100
acctccatac tttttttcca tggtggctgt attacttaca ttcccaccaa cagcatatgg
35160
tcatctcctt tctccacatc cttgccagaa tttgttatat tttgtctttt tgataatagc
35220
cattctgact ggggtaagat gatatatcac tgtagttttg atttgcattt cccttataat
35280
tagtgatgtt gagcattttt ttatatacct gttggccatt tatatgtctt cttttgagaa
35340
atgtctattc aggtcttctg cccattttta agtggattat ttgttttttt gctactgagt
35400
tcttcgagtt tcttatatat tctgatacac agccatcttc ttatgaggac tccagttata
35460
tacgattaga gaggtccacc ctttttcaga atgaaattat agcttaacta attacatctg
35520
tagtaactct atttccaagt aaggtcacat tctgaggtac aagggtttag gacttcaaca
35580
tatgaattcc agtgggacac agctcaacac atgacaccat ggtagggaac tttattctac
35640
ttgcaagttc tgagtgtctt acgcaggtag atggactggt gtgatgtatg ctttaaagac
35700
cgctgtgtga agatggcctt agggtgatga ggatggaagt tggagactaa taaaggacta
35760
agaaaatgct aagaaaatcc aggtgagagg tgatgatggc agaactaagg tgatagcagt
35820
agagagaaga gaagtggatg gagattagac atcttttgca gaacgaatga caaaataccc
35880
ctatggattg gacatgggat gaggaaaagg aaggacttga gggtggtgtc taggcttttt
35940
actttaatcg tgaagggaag ctggtgccat ttaccttgtt cggacaaacc tggagaggat
36000
caggttaggg actgcgagtg gtatggacgg caaaggaatg ggaagaatgc agggattaaa
36060
aattggaaat ccccctcccc agtcaacaat atcttacttt tatctgaaaa atactaagta
36120
aaaaagcatc cttttgttgg aaagctcaat ccttgttaaa atgaagacat ctctgggaga
36180
ggaaacatag tgagcacctt tcccaaaagc agccactgat ttggagatga gacagagtag
36240
catacaggac atcagagaga acatgctcag gacagaaaga gcaatgtagg acaaggcagt
36300
gtcttggcat cacagtcttt cctccgactg gctgtgagca agtgctcaat ttaattccat
36360
ctcagtgctg ggtcaggaca agtgcccaaa agcaaattga caaaagtacc agcatgatgg
36420
agttagaagg tagcaagttc cctccacaga gcccagctgg aaaggaagat agaggggaag
36480

ttgaccccctg gggatgggga atagggtgag aggagaacat gaaactgaga aaagggcttt
36540
gagtgaaatc taggctaaaa gctaaggttt ctttagaaac ccaccattga cccaacatga
36600
ccagggcttt ctcttgactt gattattttt gataccccat cttcttctgt attcctggaa
36660
ctagctctcc caagccccag aattgtgctt ctatcagagc tgggttttca tcagagtctc
36720
ccctttatcc tgtatctctg ttgccctatt ttgtttgaat tcctgccagg tcagctgaat
36780
ttgggcattt ggggtgaaaa accatcaagt gtggcatcct ggctttggca cctggcacag
36840
tgtgacccca ctggtctctc cctcacattt gctgtggtcc gtgcacggaa tttgtcaaaa
36900
gacctcctca gtatcagctt tcctgcagcc tcaatgcacc ttgttctgaa taggatatta
36960
ccccccaaga gtatattagg gcattttcct atgccagaag ggtccttag gcctcttgca
37020
gttttttctg ggtgacagtg aaggaggagg tggctgcaga gcttactgcc tgtggactga
37080
ccaccccagg gcctggtgtc aggaccattt gtccagcctg ttgagtgaag gtcattctgc
37140
ctaaactgta agcacaagag agagttcagc atcatttgca tcctatttta ttgtctttct
37200
tctcttttct ttcaaggcct catttttttt ggcttgaaca aatggtaaag gccattttat
37260
tacaggtacc aagccaaact ttccttggtt ttgtggccat cctgctgggg aaggaagtac
37320
tcctttactt taaataactt taaaaacatc tgtttggtct caggggctgc agctggaaag
37380
attttctaac taatacttgt tttatggggg tgtttttggg ggggtttatt gagtgtcaaa
37440
cctggcagta aattagaatc agaagacaac agttagtgat aagcagagaa gccaaggatg
37500
ttaccatagg caggcagcag agagagggga attggtggct ggccccccaa aaacagattt
37560
gaagatctcc ttctgtcatg tagtgaatcc ccaagtgcct agggtgggct gtgattactt
37620
gagctcctgt ctccactgtc tcagctcact tgccttgggg tggacacaca acacacattt
37680
gctcatagca tcaggtattc aggagcaaag agctgaattt atctggttaa tttagatacc
37740
cctacccccct cttttaacac cagattgcca ggatcatgac ctcaaaaggc taccctgaaa
37800
tgcaattgac aaatgggatg aaagatttcc cgtttcatcc acatttgcct cctgagctac
37860
ttacagcagc aggtcaccgc agccagagcc cacctgcttg cccaccatgc ccgcacacag
37920
acaatgctgc ttctgtggct ggaggtcgga acacctcagc actatctcag tttggctgca
37980
gatcctctgt gtgcttggta aacaggtttc ctcatctgta aaatgaattg gctcttccac
38040
aactttttta aaagcactaa catattagga ctctcactaa atactcaaat gctaaactca
38100
aatactaaaa gagtgcaaag ggatgggctc ccaaatatta cagtgaaggc tgcagcattt
38160
tctgaccttg ctgctttttc tggtgagtgg cttttatttc ttagtttggt ttcttctctc
38220
ccattctaat caagcaagaa gtgaccacca aaaggggcac tcaccaaacc agaacaagct
38280

agttctttca tctttaattc attgcaacca aacagatgcc acagaaagag ccaagggctc
38340
caggctttag ctccagcctt gccattaact acatatgtaa gtcagccatg ctggtctgca
38400
ggttcttgct ttgcatgatc aagggacaac ttggaaggtc tccaatcact ctattccccc
38460
agatggaaat gtattcactt atttcctgga gatgtctgtc ctcctcccag ttaaagacag
38520
accttgaccc acctccactt ccttctctgt ggccctgtct tatctgtcct cttgttcttg
38580
cctcttcaat tgttctctca ccgtgtttgt cacttctgag ctatcactgt gatccccctg
38640
attgtttttc taatgtccct gaacttcaac ctgattttca cgcatataca atgtcttcct
38700
aaacacttat agactctgac acattctgta actgacacat ttccctttat caaatgcaat
38760
ctaagaagct cacagtttct ctcagtttca acaagagaaa tcaggagcac ttgaattata
38820
caacttgaca ttattagggc tgatgtctga ttttgtcctg tctgcccctg tcatttctgt
38880
actacctttt acaaaacctc tcctatgacc tgtgtcctcc tccagctcca tttgagaaca
38940
cctgctgtat accctgtggg ctagctttta ttatgttcgc ctcaatgatg aagaaacagg
39000
cttggaagtt aaattatcta ccccaggccc acagcctgga acctaggatt ccaaccaaac
39060
cttgtctgat tctaaagcat agcagaggct ccatactctg cctccctctt ctacatcatt
39120
tcagtttctt cactttccca cctccaattc tcacccaaac tgaatgtctc acagtctctg
39180
tgcccccact ttgctccatc ccttggcctt ctgcagtcca agctccattc tgagatcatc
39240
caaggcttct cttctgtgtt gatccttggc cttcttggag tctctttctc ccatgttctc
39300
cacaacagag cattctcctg actgttttca ttctgcatct cactctttca tcagtatctt
39360
tttctctacc atgccccata aatttgggtg ctcctgaggg tcctgtcctt gtcccctgct
39420
ttcttgttgt acaacctcct tgatctactt catctactca agtttggtcc acaatttcta
39480
tattgtgaag attcaaatct gcatctctag ccatatatcc atttgcctgc taggcatttc
39540
tacctgaata ttttataggc atgccagtgg ctcttactct atggctctta ctctaagtct
39600
agactacagc agaaagcaat gctcttttta ttaaggcata gtgcctcttt cagaataatt
39660
tacagcatac aaccaggcct gctgtgcagc attacaattt gtcattaaaa ctccattcct
39720
cttgccagag taaatgagcc atttacagcc agggcgccaa gatggactgt tgttattttt
39780
tctgcctttg tattatgagt attcatggct ctcctcagac aagctcctgg ggattcccag
39840
tggagttgcc ttaacatgca ggtcaattag ccaggctcaa gggtagtttc ctggatattg
39900
gtatccccct tgcagaggac tgcaggaaag ctgaacagtg ttcccccaat gtgggtggtg
39960
atcctgagaa atatcatttg tatctgcatg tgctgtctca cacacactag ctcacatgtg
40020
cacacacacg tgcatgcaca ggacaaaacc aaacacaggg caacccagca tctgcccccc
40080

agccatcagc attgttacac ctttataggg ggcgggaaca ggttggtcag caggtgaacg
40140
tcaggtgagt tgagaaaagt tattaaatct taaatcctta aggaaagtta ttaaatctct
40200
tctaaaatgc atgcataggc gggctcagta actaacatgc aaatgtttag ggtctgaagc
40260
tcctaccgat aatctttcag atctcagaat tccagcccct tgtgctgttc tgggttgtct
40320
gacacagacg aagcagagaa cagtagaata aacagctcag taaacaattc attgagggaa
40380
agagagtgag aagattcact ggacagctag aggaggaaat actgctggtg actatggaag
40440
aaatttgccc taaggcctgc aggcaatagc ttggtcttat ttatcctggt gtcccacccct
40500
ctcctccaac acatactgcc ctggcaggta cgtagaagat gcgtgaaaat atcttttgaa
40560
ttgagctatg caaaaaatac tggattctgc cctccaagag tttactgttt agtttcacag
40620
aaagcacatg ccctcctttc tctgcctctt gaagactgac ctatctttca aggccactgg
40680
cccaattctg ttttctaagt aagaccactg agtcagtggt gacctctcct tctccctaac
40740
aaagtctgat ttacttgaat atacaactat ctccctcttg gcctgtgaat ttcttgtgtt
40800
agggaacata tctgatttat ccttatctct tccacagtac ctggtgtaaa atgcccaata
40860
aatgcattga aatattcatg aagcttacta aatgctctgc cttatgagcc atgaaatata
40920
aagtgcctta aactttgttt ttctcttatg taaaataagg ataataataa tgacacccct
40980
ataggattgc tgcaaggatt aagtgtgata atatatataa aactcttagc acaaacacct
41040
ggctcacagg aatagtagct actaccataa tggtaacttc gagggcaagt tttctcagag
41100
ttatttagcc ctccttcacc ctgtgtccag gagtgcagat cagaatggtc agattccagg
41160
acaccaagtt ttctgtggga gcttccctag gaatataact aaggaattta aatcaggttc
41220
agctcatgct gttacactct cttcctccac tcaggcattg ggtgtggctt ttccaagctt
41280
gagaagggtg tgatctgaga tgggcttggg tatagagggg aattatattt aggtctaccc
41340
tgtataggaa aaagtgcctt cccaaagtct ccctggccta aagtataaga gatatgtgtt
41400
gggatttaga cccagagccc aagccaataa tgggaccccc ttctcacatg tggctacctc
41460
ctgctatcac cacaacagct atcataccca taactacaac agaggccaat taacgtggtg
41520
ataattgaca aatgtcaaga catcctacat tgaggcacac tgtgcgtttt gcgtgagctt
41580
ttaaattggt agggaaggaa aacttttata cctacaccta tcatggaagg cagaaggtaa
41640
gagctaaaat aaaggtatgc caagaacaaa ggcaggaaag aagggtttta acaacttgag
41700
gcctgatcca ttgattagtg aagaggaaac atgttcaaaa accactctat aaccaccttc
41760
tccaagtttt ttataatttt gcttcttcgg atatcttctc atcatagtct taaatgccat
41820
caaattaact gaaaaatgct aaaaatgcaa ccactctaag agaatgggtt agatgggaga
41880

ataacaaata agaaaggatg ccaccttctt ccaaccctca tccctcttcc ttttgacagt
45540
tgcagatgtt gctaattcat tttggcaccc tttttctctg acccaaatat agtcttataa
45600
accttttcaa cccacggctc taggcaagta tcaccttttg ctcttttggc accagatctc
45660
ttgaacacta tttactggtt ttggaaagat tatacatgta tgtctggagt tgaatgactg
45720
aacagagcaa taataagagt taaagcaaga aagacaggcc tacaggagat ggcagagggt
45780
cttgcctgtc aggcattgat tttgaacttc attgcatagg caatcaagaa ctattgaagt
45840
ttttgcacaa aagactatag atgagattaa cctggttacc gtaaaggaca aagtgattgc
45900
aggtagaatg aggccagctt cataaatgaa tcatcaggat atgagaagca agggcttgaa
45960
catgagaggc catagtggga atggagggaa agggacaatg tgagaagcag tgaaggagaa
46020
gggctgattg agtaaagcag tggagaagac agtgaaagat gtcagatgac taccatgttt
46080
ggcgactgag tgagggaaga ggtggtgatg atattactga agagagaggc aaggggtggt
46140
cactggattt agagcagaca ttatcaactt gtggtgtcca gacatttcac cctgggagaa
46200
acctgttctg aagtggcttc agcatctctg aggtcagatt cctagttcta ctatttttct
46260
actgactgaa atggaaatcg agtaggcaag gcttttgatt tgtctcagtg gtctcttctg
46320
taaaatgggg gtgtttatat ccatagtctt atcacagggc tatttggggg attaagtaag
46380
acaagtgtgg cagagctttg taaactgtaa tacactgtgt acaattggat aattatggat
46440
tcttctgact catccacatg gatgtctgct gaccctgggg gaccggagcc tgggagggag
46500
gccagacctg gaaatggaaa cttgaaaatg ttctctgtag aaaagataat taacatttga
46560
ggatggttaa gtcctcttaa atagatgtca gaaaaaatgg aggtcatgta gacagaatgt
46620
tggataacac tactttgtaa aatattttat cttatttcca ttataaaaga aaaaaagctg
46680
ggctgggcac ggtggctcac gcctgtaatc ccagcacttt gggagactga ggcgggtgga
46740
ttacctgagg tcgggagttc aagaccagcc tggccaacgt ggtgaaaccc tgtctctact
46800
gaaaatagaa aaattagccg ggtgtggtga caggtgcctg taatcctagc tactcgggag
46860
gctgaggccg gagaattgct tgaacccagg aggtggaggt tgcagtgagc caagattgca
46920
ccattgcact ccagcctggg cgacaagagt gaaactccat ctcagaaaaa aaaaaaaaat
46980
agacaggaaa ataaaaaaag ccacctcaca tagtctacta ccaccaaaca catcattaac
47040
attatatttc tttattccat gctctttgtt tttaatataa acaattactt ttaagggaaa
47100
atgagaaaag gagagagtga taagacttta ttttaaaagg tggaataatt ctaaccatgg
47160
agagtattta taaatttttt tttttgaga cagagtctcg ctctgtcacc cagggtggag
47220
tgcaatggcg tgatctcagc tcactgcaac ctccacctcc cgggttcaag caattctcct
47280

gcctcagcct cctgagtagc tgggattaca.ggcaaccgcc accatgccct gccaattttt
47340
tttttttttt tttttttgga gatggagtct tgctgtgtcg ccccaggctg gagtgcagtg
47400
gcatgatctt ggctcactgc aagctccgcc tcctgggttc acgccattct cctgcctcag
47460
cctcccaagt agctgggact acaggcgccc gccaccgcac ccagctaatt tttgtatttt
47520
tagtagagac agggtttcat tatgttggcc aggctggtct tcaactcctg gcctcaagca
47580
atcctcctgc ctcagcctcc caaagtgctg gaattacagg tgtgagccac cgtgccaggc
47640
ccataaaata ttttttatag acaagtgaga gcagaaatca caggttctta tgagcaggaa
47700
aattttgaag gtcatctact ctgaacgttt ttttgtttgt ttgtttgttg ttgttgtttg
47760
tttgtttttg cttagtttac atttattaaa tacccgttat ggtccaggcc cttggctaag
47820
cgccatccat gcaatatatc acaagatatg cccagcaatc ctaggaggta gggtttatta
47880
ctacccatcg tacagaggag gaaactgagt catagagttt tagtgtcctg atcctggtca
47940
cagagccagg aagtggcaga gcaggccagg ccaagtctgt ctgacatcag agctcatcag
48000
agccctcccc attgtccttg aaccagtaaa gatggagttc ttctacaggg gtggttgggg
48060
gacaaggacc ccatgggtgt gtctgagtca gaaacatctg cgagtgggct gagaaatgag
48120
tcttctgtga aaaagagcaa aagaaaaaat gggtcaggag ccaataatca ttgtccatct
48180
ttgtgtgaat gtatggtgtg ggagtgggag caataaacga ttctaaggtc acacagaaaa
48240
gatgccacct tctccaatca cataccgccc ctcgtccccc agttttctct gaaatagctc
48300
ttcttttggc tctatcctgg cttcttcaca caggggtgtc cagtcatctc atcctggtgg
48360
gacagggata gagctgtggc agtggagatg aggaagctcg cctcctaagt gagtctgaat
48420
tcttaaatat ggagccactc cataatcatt tggagtgaat attgggccat ggcccttttt
48480
cttgccagct gagctatgaa aaaaggatgt cctaagacca gaggctgtgg gaccattccc
48540
agcccctgca ggaatcaaag gagctgacag aattgtttgt ttgttttttt cacaaattga
48600
aaaaaaaaat gtaaaatttt tgaaaagaaa gcctcattga aaagaaatcc ctctccccag
48660
ctgggctccc aggcagcctc ctgcagaaca tccttagcat tgcagagttg ttcccatggc
48720
aaccgagtaa ggggcttttt gttttcctta gaagattgaa tcctttcaac cagaaggtaa
48780
ccactggttc ttccccacaa tccacactcc aaaccccccta cccttatttg actacatgac
48840
tagttttgca tttatggatt tttttatgcc taattgaaaa aggctaaata tacagaaact
48900
gaggctgaag tggtttaagg aggcaactgg cccagtggtt tctcagcaac cacatgtcaa
48960
agctgtggac gttagacttg acgagagcaa gacatatcag aatctgtagc aggagcatct
49020
agtctcccag ttcaatagtg tccacaaaag aaatccagag gtttttgaag caaggaattt
49080

gggtggcact gctgtgagaa acaatcacct ggctcctcca tggggcatag agtggagatg
49140
cttcttcaaa taccccttcc tttccaaggc catgactcag aatgactggc gtagggagcc
49200
tggacctgat ctcttcaagg aaggggaatc agatgagctg tttaatctct cttgtaaaat
49260
gaggggttat gagaccatag gctcattttg ggggggtct aaaatgcagt attttttgaa
49320
ctgatatggg gaaaaaaaga catttctgaa ttgttgtcat gttgcagatt ctgggccgtt
49380
ccagcataag cacctttctt agagtacttg gctttgtgaa gtagtcctta tcccctcctt
49440
ccactatttt acatcaagtt aaaatagagg aagatgccta gaaatggccg tatagacaga
49500
gaaaactgca ctaaaactcc ctccgtcatg cctgactcct ctctagacta tgaccatcga
49560
ggggccagaa atcatatctt aaagatcact gtgcctccag tacccagcac ggtgtttaat
49620
aaatgtttgt tgaatgaacg aactagtaaa attttcaaat cattagagct gaagtatcct
49680
ttaagattct ttagtccctc attttacaga taaggaagct aaggctcaag acattgtgtg
49740
gcttggccaa aggcacacag caagctaaag gcagagggag gacaggaccc ggctgtctca
49800
accccctggc tgctacactt cctgcagcat ttctaattct tttaccattc ttgcgaggga
49860
ttttacaggc atgtactgct agagccgaaa taattagaag cctcttacta ctcatcagaa
49920
aagctatgtg agcccctagg gaggacacag ctagcctaga ctctgcctct ttgccctctg
49980
ctgcttatta gcagaatgta agtggttgtg tatgatgatt agtgtaagta ggatgggcaa
50040
atgcacacct ttcccacctt caaactcaga agttgtaacc aagagtcaca ctgactaaac
50100
actccaattt ccctttctgt ttttcttaac atatgtccta ttttaccaat aatagccatg
50160
gtatattagt catggtattt cacgctagct gcagaaataa cttccaaatc tcattggctt
50220
actcagtgaa agtttatttc ttactcatat aaagttgaat gtcctggtca ggcagttatc
50280
taagccacaa cttggggatg gggatgcagg cagcttccat cgtattggct ccaccattca
50340
gggatggcag agttgctctg gcataatcca accaatagag gggggaggtt tggcacttgt
50400
cagttaacca cctagcctag cattgacaca caccacttct acatacactc ccctagtcat
50460
cattcagtca tgtggcccaa cctagatgca aaggcatctg ggaaatgtag cccctatctg
50520
gtcagcaaca actttgcact tggaagggga gcctgaatcg ttattggtct ccaacacatg
50580
taactagcaa ttatacagaa cgttatttgt caggcaatgt gccaagaatt atttcattta
50640
atcttcacaa caatcctatg aggttattgt cctctttaac gtatagatga aaaagttgat
50700
ggtagagata taacttaact aatgcaaagt tgcataagtg gttggtagca aatccaaaat
50760
tcaggctgtt ctctccagag ctcaggctca tgattgctgc attctactgc tttgagcttc
50820
tgatctgaga aaatgcatca gccactaagt agcctgtgta gtctccagca attactttcc
50880

tccctctgga tcttggtttc attctctgca aagtgaggat gtttaactgg ataaaatctg 50940
atgtcacctg ccagctggga catcatatga ttctcagggt aagcatatca ggtgggtggg 51000
gtccccagtg atgcttgacc atagcaaagc cctttcaaag gtttcttagc acaccacata 51060
aatggaagcc tcacagtgtc catgtaggag aaagcagggc aaagtatttc catttaccca 51120
acaaagaaat caacatatag taaaaagaga gtgttttccc accaaggcct cagattgact 51180
agcggtagcc ttggaaatag gactttattt tgtatagtac ttttgccacc agggtggggg 51240
ggaaaagagt gcttctttgc cccaaatgct ggtttcataa aacctaaaga tgtcacatgg 51300
aaacacacca ttcccccaat ccccctcaaa aaactacttg cacttaaatg aaagagtaaa 51360
gctgtaggac tttactgagc agtgtcctgt ggggtccttg cactgccatg ctcttgaggg 51420
gctcgaggtg tatgaattcc ccagcattac ttctccttag aggtttcaga tgagcagtat 51480
gagctccaaa ctcatgctag acccaagtat ttcatgaaag aacaatcctt gaatgacttt 51540
atacagcaaa gctatatttc actgtgtcct agaaaaccaa ttgtgtgtgt ttgtgtgtgt 51600
gtgtacaact gcttgtgttc tttctaccta tgtccccctg atgcctccac acagaacatc 51660
ccaaactcca tttcaggttc ctcttgagat tccccaaactt ggaaacagga gatgcttcaa 51720
aggcctcttg gaatgtcttt tgaggcttta tattgtgata tgttggacag atggttaaga 51780
aacagaagaa gagcatcacc aaaaggattt ctcattttat gtggagatct attaatattt 51840
gccactagca aaggcattct ttcttgggaa tgaattatgc cccctagaatc agattgaccc 51900
cacagaaaca agggagaata aatagagact tgagcttaga ccttacaaca tggccagagc 51960
tgaaaaggct gagctctagg cagagaagat gcaagagcag cttcagaaga cctgagagct 52020
tatttgggta ggttcctctg gtgtaaaggg ttcttgttca cgttttcttc cagaataaga 52080
aaagaacgca aggtgtcaga gggtggatgg aaacagggta taaagcagga gcatttggaa 52140
tctgcccttt gtagcctggc ccagagagcg tcaggcagct tgttgggtaa taagtaacac 52200
tggcattttt cccatggttc tgtcatctta aagagcagga tacataaagg gattcagatg 52260
tcttgttggt ttggagaagc ttctttttaa taccttgttt taaaatttac ctggaattta 52320
ttttaatcag gtgtggtaag atgcacagac atggagatga cagtcatgaa ggaagaagta 52380
tttatactca cagatccctg taaataggaa gcatggcctc catgcaggcc aatggggaag 52440
caccagggtc agccgcaagg cagaaggagc aagaggaaaa catggacaag aggctctact 52500
gtggattcag tggcaaagaa tgggaggggc agagtaagca ggtttaggat tatcgggttt 52560
52560
gaatgacttg attgagctgt agggtgtaga gactgcctct actgtctggc accaggggta 52620
attagggcag ctggatagtg gtctggagtg tgagagctcc ctaaaggagg tggttggagg 52680

tgtaggtttt ggattggttg atctgtatat gaaaggtgca cgtgcaggtt gagtcctcta 52740
ctatcactag aaattggctg gtcccaggag aagtagtctc tctagagaca gcaatgcccc 52800
agatgtcaaa gcatcagaaa atacagaaaa aaaattaaaa gcatgattaa ttcatactca 52860
caggtctagt ttttgtgtag ttaagagcaa cctaaagaag ttgataactc gtgttgcagg 52920
tcaggtttcc cagaaatcat attctcagat gaagatttgc atgaaggagg tttaatgctc 52980
aaactaagcc ctaaggctcc atacctgtgg aggaagtgaa agaagcccaa ctgggcacag 53040
aaggtggaac acaatgccag tcacacaaag acctcagtgg atcctgggcc atgaggagct 53100
ctaagcacag atgacccttc agaaatgtct ccaagtgggg aaaggaatca tgctagtcac 53160
tggatgtggg cttcccactc caccccatga gggcatgacc ttaagtgaga gagctctttg 53220
gacacagggc atcctaagag gggcactcag cagccacatt gggcaccaag actctcagca 53280
gctagaagaa gaaggtatag tcccaaaggg gaatctgggc tgcacacctt agtatccatt 53340
agaactggaa gtaggctgaa tcccaggcag ggatcccctg gagaacacag gtaatttttt 53400
aaaaaatcaa gctatgtgtc tgaggctatg tggtaagaca tctcagtttt ctgctaggaa 53460
aagccaccaa accagattgg cttattcatg ttgaaaagtc tgagaatcac actcagatgt 53520
tgttgataat tctgcttgga taaaatttat ctattggtat gcttgtgata tagcagtacc 53580
attgctaaaa attccatgcg gagaatccaa tctgcatcat tttctttctc aatgatttgt 53640
ttttaaaggc agaggttcgg ctgtgcccct ttaaaccttc tgtgcaagtg ccagcttcct 53700
ttcaaatgga gaagcagcag ccctgtcaga aagggtggct ggagctcccc ttttgtgaga 53760
ggaggaaaac ttactgggaa ttacctgttc gagagccaca catgaaggca taccactgct 53820
tcctctgacc ttccagccgg tatattaatg acatactgtt gtacctgaga accaatgatg 53880
aagtgggtga tgtgcctggc accttaaagg cctgggcctg ctttgacagg ggagatgata 53940
cacaacatgg ctgttagcca gctctcactg catctggaag caccatgttc cttagagcca 54000
aagttctcaa actgtgcttc ctgctgggct ccacagatcc ttcccgttcc accctgcaca 54060
caaacgtgca cacacataca cacacacaca cacacacaca cacacacagt gttctcaatg 54120
ctcgccattt agttagtatg caccaaatat gtgtagtatc tggttccacc cctggcctct 54180
cagacaatta ttagtatttt tgggagcggg gaggagagtc aggaagaccc aagcgccata 54240
tttattattt ccccagccac cccggcccag gctacatcca agttcaaagt ctatgacccc 54300
ctctctgagc tttcagcact acctcccttt gtgggggagg ggggtgccaa ttctctttct 54360
tctcatcatc tcctgttgca aaataaaagc ctaggcattc ctttgagaaa cttgggcctg 54420
gcattggaag gcgtctgaca aaggctttgt taaatgagtg gagggaggga cggtctggga 54480

gatacttttt caggtggcat aggacctccg cttcttccct tctcacatga gaaggaagat
54540
ttttctagaa atctacaggt gtttaagctg gaatgtgcct cagacatcat ctggttggac
54600
cctttcattt tgcagatctg aggcctagaa agatttggta acttgcccca ggtcacagtt
54660
gacagaattg ctcagtgaaa agtccagcat aaatacccca gcccatgtgg ccactggctg
54720
tgtgctcagc tagtgaggca cacttacttc ttaatttgtg ccacccactt ttcaggctcc
54780
cttaggacag cctccacctg ctcctactgt gcttcccatc gtccctctcc tcaggcacag
54840
gctgaggagt aataagagca cctgatatgt gtcaggcctt actgtgtgct aggaattgtg
54900
ctaagtactt cctatgaatt ttccatttat tctttataat aactttgtaa agttagagcc
54960
attattccag aagggaaaac cgaggcaatg ggagtcaaag caaagaattt gggcttttaa
55020
ccattacact attttgcaca agtagccagt aatgaaaagg ctgctatccg gaatcatctt
55080
tgcaaaaggt aatttcttta gcactttatc agaagaaggg ggctccttcc tcaaattctg
55140
agggaagaga agtggggaag aaaagatgac tgaatccaaa gctcgggcag ggaaagcaca
55200
tcgagtgcca agtgcgctgc gctggggtct agtcctgact cagccgccat cttcccaagt
55260
gcttcctgga attctctcct ctcgtggggc ctcagctcct tcatcttagg aaagaagggt
55320
aaagatctac agacaaattg atctttaagt atccttagag cactaccatt ttcagaatct
55380
aggattctat atccttccaa ttatctctgt gtagggaatt attggtcgtg tctcctgatt
55440
agggagccgg acactcgtct gtcagcccca cctggctctg caaagtccct tgtgtatctg
55500
ccctgcctgg tcacgggaga ggaagagaca aggaaacacc accgctccga ctctgtggag
55560
cacgcgctct ctcccaccca cacacccgct caggagagga ggaacctgca catttgagtc
55620
tcctcagagc ctctgcagac tcccagcagg ggtctggctt tcctctcagg tagcacagtc
55680
atgctgtaaa ctcatttggg tcttgcttgg tatgataatg cgtttagttg aagggttata
55740
taattgcaga gtcgatgatg atctctaggc caatttaaag tcaaagctat ttttaatgga
55800
attgccagag gagggcaggg atgggggcag ggaggagaga tggttagaga gtgcttttga
55860
aaccaacctc caacaatttc agccattgca tttccgaacc tgaattttca gggcagaaat
55920
tggacaatgc caattaaatc agagcaggtg tatgtgagag ctgggttcac cttcttgcag
55980
ctacagtttt attttgaata ctgttgcagg tagtgaaaat atgactaggc tgaataagag
56040
atctcagtct attcccagct cagccaaaag cccttagtgt gtccttgatc aagttacttc
56100
ccctatccat ttccttacct gcaaatgaga agcttgaacc aaactatcct aatgtccctt
56160
tcaactctaa aatcctagat gatcctcaga tgtcaacagt gctgaagccc agcactgtaa
56220
gatgtcaggt ggtccgcaga gggtgaggct cttcctgctc aaattatttc ttccacccaa
56280

gactcctcag ttacctctgt acacaacctt gcaggcccat ctaagtatcc aataacctgg
56340
ggctttagtt tacaaatttt cttggggaag aaggtaaaag ggatctagct ttctgggtta
56400
tgaatgccat gtagggaggg catggtttga gttagtcctg gtgctgggag ttcatgagac
56460
ttattctcaa atcttcagag aagaaaattc cgtgaacacc tgggaacatc aggaaaaaaa
56520
aaatgtcccc taggctactg tcaggttagg ctgctggttc tgatttgacc ttgaacttgc
56580
tataattgaa caagataagc atgtgaccta atgaaatact ttaaaacttg tagcttcctt
56640
cagcacagaa gtggctctct gaaccaattt taagcaatcc tggctctatc tgtgcatgtt
56700
gatttagcct gtggttatag tgttaacaat ttagtgattc acctcatttt taatctctct
56760
ttccctttag caggatcatt ttctctgtgt taagggatca acattgaggt aagaatggct
56820
aaataatagc atcttctgga atacaaatga ctttataaat aaaagaagat aaaaggaaga
56880
agtaggatga tttctcagct ctaatacact tagcaaatgc catatgcttt ctcctgcgtg
56940
tactggtcag gccagttcta gatacaatca tgcgctgcat aatgatgttt tggtcaacag
57000
tggattgcat atgtgacggt agtcctttaa gattataata ccatattttt gctgtgcctt
57060
ttctaggtct agatatgttt agatacacac atacttacca ttgtgttcca attgcctaca
57120
gtttccagta cagtaacctg ttgtacaggt ttgtaaccta ggagcaatag gctataccat
57180
acagcctagg tgtgtagtag gctataccac ttaggtctgg gtaagtacac tctatgatgt
57240
tttcacagtg atgaaacttc ctaatgacaa atttctcaga atgtatccca gttgttaagt
57300
gaggcatgac agtactatat ctcaagactg tccccaagct gaagtctcca gtggacacaa
57360
agaccaatgt atttagttga atcgtggacc ccaaaagttc aagtccaccc agaacctcag
57420
aatacaagtt caagtccacc cagaacctca gaatacaatt ttatttagaa atagggtctt
57480
tgcaaatgta gtaagttaag atgaggtcat accagagtaa agtgggccct aaatccaata
57540
tgactagcat ccttgtaaga aaaggaaaag gaacacagac aggggagaag gccatgtgag
57600
aacagagaca aagactggag tgaggcatct acaagacagg gaacaccaag gattgccagg
57660
agccaccaga agctaggaag aagcaaggaa gcatcctctt ctggggcctt cagagacagg
57720
atggccctgc tgacaccatt gtttcaaatg tttagccttc agaactgtga gacaataaat
57780
gtatattgtt tcaaaccatc cagttggtgg tactttgtta taggaaacta atacattcag
57840
gatggagagg tgtctgggaa gcccatgaga acaaatggaa agagccagaa gccctcaacc
57900
ttggctcgtc tacagcccat tttcttcatt cccgcatcca ggctttgaga tgacaggaag
57960
ctgtgaaacc tgtgaattgt ctccaccgca aatcctgctc cctggtccca cctagactgt
58020
cagggttgtg tggcaaggct ttcatgcctc tcactgactg cctagtacgt cccctcaatg
58080

actggtccac atctttctca cctttctcat gcatggcccc agatccaccc cagtgcctcg
58140
tcctcaagag gtgatttatt ccgagacact gatgagagca ctgtccttcc tgtgtctgag
58200
ggaaggcatg taactcttgc ttatcttcac ctgtgctcta gatcctgacc ttctctggca
58260
acctcaggga ccttgcacca tccattcttc tcgcctaatg gcgagactca gtctctccct
58320
ctccctttcc actctccctt gccattctta gtatctttct acaagcaggt cttccaaagt
58380
actgcttgag gtctgagttg gagggaacat gcctctaccc tactaaaaag agaaattcct
58440
ctgcagaaga cccaagctga ctgacaaatc cctttactgc aactgcagct ctagctccca
58500
ccattttcct gtacttactc tcctgctcag gttccctggc attgctgatg tctttcagcc
58560
tttgtgccct ggccccttcc ctcctctccc ctcatctagc actacctgtc aaaatcaggg
58620
acttacttta aaatttatcc caaattatca ttgccatcat ctccactgtc accttatcat
58680
atgtttgaat agcgtttcca tttcccaaat gttttcgcat gcactttctc aattgagcct
58740
tacgaatcct agagctgaga agggtaacaa tttatgagtc ctttgacaaa tgtggaaact
58800
gacatcacag aaagtaagtt gccagccgat atgtcactgt cttcaaactc ttctttgtat
58860
ttttattatc tcccattata ttctgcctct tgtaatgatt atttctacat tggtcatatc
58920
tttccttctg tactgatctt cgcttatgat aacaaataat aatagtttac ctttgcatca
58980
cacttgatgg tttacaaaat gcttcaaatt caacatggcc cttgatcctg aagatattta
59040
tcacttaaga atcattatcg ccattttaaa atacaaattt attacttggg ctaaattttc
59100
ttattatagt tgggataggc cttcatccat agggtgagtg cagtatttgt ggactgtcat
59160
ggcagcttaa acatttagta cttgaaaatc tgatgcattg atcatcagag aaatgcaaat
59220
caaaactaca atgagatatt atttcacccc agttaaaatg gcttttagcc aaaagacagg
59280
caataatgaa tgctgacgag ggtgtgaaga aaacggagct ttcatacact gttggtgagg
59340
atgtaaatta gtacaaccac cagggaaaac agtttggagg ttcctcaaaa aactaaaaat
59400
tgagctaccg tgtgatccac caatcccact gctgggtatg tacccaaaag agaggaaatc
59460
agtatatgaa agaggtatct gcagccgggc gcggtggctc acgcctgtaa tcccagcact
59520
ttgggaggcc gaggcaggca gatcatgagg tcaggagatc gagaccatct tggctaacac
59580
ggtaaaaccc cgtctctact aaaaatacaa aaaattagcc aggcgcggtg gcgggcacct
59640
gtatttccag ctactcggaa ggctgaggca ggagaatggc atgaacctgg gaggcgtaac
59700
tttcagtgag ccgagatagc accactgcag tctggcctgg gcgaaagagc gagactctgt
59760
ctcaaaaaaa aaaaaaaaaa aaagaaagag gtatctgcac tctcatgttt gcagcagcac
59820
tgttcacaat agctaagatt tggaagcaac ctaagtgccc atcaacagat gaatggataa
59880

agaaaatgtg gtacatatat acaatggagt actattcaat aaaaaaaaag aatgagatcc
59940
agtcattagc aacaacatgg atggaactgg agatcattgt gttaagtgaa ataagccagg
60000
cacagaaaga aaaacatctt atgttcttac ttatttgtgg gatctaaaaa gcaaaacagt
60060
tgaacctatg gacatagaga gtagaaggat ggttaccaga ggctgggaag ggtggtgggg
60120
ggcttagggg gagggtggga tggttaactg gtacaaaaac agaaagaatg aataaggcct
60180
actatttgat agcacatcag ggtgactata gtaaataata acgtagctgt acatttttaa
60240
aaaacttgag tataactaaa ttgtttgcaa ctcaatggac aaatgcttga ggggatgaat
60300
atgccattat tcatgatgtg cttatttcac attgcatgcc tctgtcaaaa catcatatgt
60360
acccaataaa tatatacaac tactacatac ccacaaaaat taaaagtaaa aaaaaaaatt
60420
aagaaaataa aagaacaaaa gtagatgtat tctacatgtc tccatattgt aaaactagaa
60480
ccagtcagtt aactttagag gaaggggatt gtggacttga tataaagaca actttataat
60540
atgcagagca gcctaatcct acaattgtca aaaagtatag tggattcttt atttatttgt
60600
ccatgatatt atagaggtca tttctgcttt aacaagtagg tgggagatag ctaggtagga
60660
tatattttgt tcttattttt tattttaaaa tattgggctg tggctggaca tggtggctga
60720
aacctgtaat ctcagcactt tgggaggctg aggcaggcag atcacctcag gttaggactt
60780
ttcgagacca gcttggccaa tatggtgaaa ccccatccct accaaaaata caaaaattag
60840
ccagttgtgg tggcatgcac tgtagtctca gctccttggg aggctgaggc aggagaattg
60900
cttgaacata ggaggtggag gttgcagtga actgagatta cgccactgca ctccagactg
60960
ggaaacagag tgagactctg ttttatatat atatatatat acacacacgt acatatacat
61020
gtatatatat acacattatt attgaaagca gccaaagaaa aataacacat tatatataga
61080
gaaagagcaa atgatgagtg actttatatg tatatatatg tgtgtgtgta tatatataat
61140
gtgtatatat atacatatat atatataggt taagaacctt cagcacatgt atacctatgt
61200
aacaaacctg catgttcagc acatgtatcc cagaacttaa agtgaaaaaa aaaaaaaaga
61260
accttctgca tgccagtaac tgtgctaagt gattaggatg caatggtaat aaaaacaaag
61320
tccctctcct taaagaattt tctatttaga agggaaaact ggtaaataaa aaataaatat
61380
ataaattaca atttgtgaaa agtgctacac atgaaagagt gctgagacag acatcaatgg
61440
ataaacttta gattgagaag ggctctgaca aagcaacatt taaggtgcaa cctgagagaa
61500
tagaagttaa acaggcagat attggtgaaa gagcagtcta ggcagaggga acatcatttg
61560
caaaggccca gggtaaagaa gatcctggta aggaaatgac agtggaagaa ggttagtgta
61620
gcaggactgt ggctagggcg gagaggcagg gaagtagttt agaatttcaa tgcaatagga
61680

aatatggaag attgaaggca gttttgcatt ataaaataat atgattgcta ttttaaagct
61740
actttatcta aggatggaag attcttaaat aaacttgtgt atacttggac cacaccacca
61800
tgagcagcag ctgctctaat tcagagcagt cctcctgcca aacgctgtgt gagacaaagc
61860
tctgattcat aaaggggcat ttttctctgg gagaaaacca gtgatccatc tgtagaagta
61920
cctgagtcta aggggagacg aagcagcaaa agaaattggc ttgtgaggac agggacattg
61980
taagaatgaa aagaggaagg gaggtgctga gcccttttc ttttttcttt ttcatttttc
62040
ttttttttt tttttgagac ggagtcttgc tttgtcgccc aggctggagt gcagtggcgt
62100
aatctcagct cagtgcaacc tccggctccc gggttaaagc gattctcctg cctcagcctc
62160
ccaagtagct gggactacag gcccttttc ttaatccaca accttcagtt ggattttgca
62220
aatgagtctg tcttcactgt ttccattcag tggctggaga caacttggaa gagaatctca
62280
gaaataactc tggctgctca cccagttgtt tgtaaatttt tattgagact ctactgtgtg
62340
ccaggctgta ccaggcactc agatatgaca gtgaatgaga taggcaacat ctttgccatt
62400
ggagagccta cactgaagtg gacatgaggg agttgaaagc aactcttata ggaaatcatg
62460
gtaaagacgt ccaagagaag aaagatgaag ggcaaacaca tgcacggatg ccaaacatct
62520
atcagagaga aaggaatttt cagacctgac ctgaatgatg aaaggaggtt tttggaaagg
62580
aaaatagaag ggaaggacaa gggaaattat ctgggcagca atatttatct gctgtggtgc
62640
ttcactctct ctctaatcct tttccacccc agccccaaat ttgaaaggat tgcagggagc
62700
tcctgctgga gtcatttctg gtattaaaaa tgtacagaaa ggaaagcttt ggttctgagt
62760
ttgcaggctt ccctgtcttt cattcctatt gtagaaagca gcttatataa aaagatgtgc
62820
tgtgtggccc tttgagctgc tgtgattgtg ttaggacccc actggatggt attcgcatga
62880
attaatctac tgtagcatct ctacaaatca agaggctggc ttctgtttga aatgtcccaa
62940
ggctttgtgc acagggcaag ctaaatgtct ccctacagtg agactgaaaa tgccttgggt
63000
gcccttgtcg ataggatctg atatatagat gcatgtctac aattgcacag tggctgctgg
63060
caacatttat tacaatctga atgtgaaatg gctattctgt tcaaggattc tgataaaaag
63120
tatcagccac agtagatgta taaggagcct ggtttcactg caactgacta cagttatctg
63180
atttttttt tctagttcat ttttagtctg tggagcaaac agagatttcc tccccaaatg
63240
atgtcctttc tcagtcacca gggtgtggtt atttggtttt atgtagagga gatagaaacc
63300
aatcagtcta aatcatattc tgttgaaatc agaaccaaag gatccacaat ctggctccaa
63360
tctaactttc cagcctcaac tcctacctgt tctttgttac tcttacccct ctaaaccact
63420
tgtgggatcc tgaacttgta acctgtgctc agactggtgc ttttgcactt ctctgatggg
63480

aaagatttct ctcatctttt atgattcagc tgaagtttca atgcttctga aattttttcc
63540
tgctcctgct ggagagcttg tttcttctgg attcccatag gtcaggtcct gtgtttggca
63600
ttgggataca aagccaagta acatagcatc catattctca aatcctcaca atttggtagg
63660
aatatagaca agtaaataca ccctgtgcaa ccttttgtaa cagaggtata aaagggtatg
63720
aaataaagaa tttaatcaaa tcaaattgaa tatgggcttc aactctgaga tcttcttcca
63780
tgatgaggtt cccagtttac tctagtgagg tcatgattcc atactggcac tcttctaggc
63840
acataaggct ctatcctatt attaaataaa gattattacc attctcactg caagcagcag
63900
caacctgaca ccatcatcat cataaaataa gtaaaacnag agttaattaa gtgtgaactt
63960
tctaaaccaa cattgtatga gataattact cataaaaatg attcttcact ttccaaaggt
64020
gcctctaaat actaagattt cagttacaat aaaacttaga tccaatttac agatattaaa
64080
tttggtccat tttccaagaa tattttcttt tctcataaaa taaaaaaagt atgtgagaat
64140
attagcacaa aggggttgca aaataaattt tatttatcca gatgtgagat aagaggcaca
64200
tgcgtctttt ttcttgtttt actgcactgg ttaggacctc tagtatgttg aataaaagtg
64260
gtaagaatgg acattcttgc tttgtttcca gtttgcttta atatgttttc tgtcagtttt
64320
tcatagatgc cttttatcag actgattaat tcagtctatt attatttcag tatgttattc
64380
agtttattat ttcataataa ttttttaaac catgaatgag tttgaatttt gtcattcctt
64440
tatgtatctg ttgaaatgat catatcgttt tgctttctaa agcttctaat atggtttaat
64500
cacatttatt gatttttcaa atgtgaagca aatttaaatt catggcataa atcctacttg
64560
gtcatcgatg tgttatcctt tttgtatgct tctgggttca atctgatact attttgttaa
64620
gtatttgtgg tgtcttttca tgagagatgt tggtctgcaa tttttttttc ttgtaaggtt
64680
tttgtaaggg tttaagaaag caaggtcagg taagcttcac aaagtaagtc aagaagtatt
64740
ttcaccttta tcttctgaaa gaatttatgc aacgttgaaa ttatttgttt cagagatggt
64800
caacagaata taccagagaa actatttgga cttagagctt ccttgggggа aggtttttga
64860
taaataatgc aatttcttta atacatagta cttatatttt ctatcttacc ttgtgacaat
64920
tctgatgaat tgtgtttttc aagaagtttg cccatgtcat ctgagttgtt aaacttacta
64980
caacaaagtc tttgataata ttcctatatt agcctttgaa tgtctataag atctgtcctg
65040
atgttccctc tctcactttt ttaaagaagt cttgctagag gtttaccaat tttattttgt
65100
tttattttat tttatttttt cttatttgag acagagtctc gctttgtcgc ccaggttgga
65160
gtgcagtggc tcgatctcgg ctcactgcaa gctctgcctc ccaggttcac gccattctcc
65220
tgcctcagcc tcccgagcag ctgggactac aggcaccagc caccatgccc ggctaatttt
65280

ttgtattttt agtagagacg gggtttcacc acgttagcca ggatggtctc gatctcctga
65340
ccttgtgatc cacctgcctc ggcctcccaa agtgctggga ttacaggcgt gagccaccgc
65400
gcctggccga ggtttaccaa gtttattaat cttttcaaag gactacattt tggctttgat
65460
aatttttcct atttttatc tacattatac tgattccaat tcttatcttt attcttttct
65520
tccttctctt cactttgggt ttaatttgtt catttttttt tctggcttct tgagatagaa
65580
gctgagatca ttgattttga acctttcttc ttttctaaat aagtgcattt aaacttacac
65640
atttcccttt aagcactgcc ttagctgtat ctcacaaatt ttgatattgt cttttcattg
65700
tcttttattc aatatattct aatttttctt gtgatttctt ctttggccca taggctgttt
65760
agaaatatgt agttagtttc caaatattcg aagactttca cagatacctt actattattg
65820
atttctaatt taattctgct acaatccaag tatatacatt ataaagtttc agccttttga
65880
aatgtattaa gaatattacc agagataaga agataagaat attaccagcg ataagtaggg
65940
atatttcata aataatagac gaattgattc atcaagaata tacaacaatc ataaatgtgt
66000
atgtgtctaa taacagagtc tcaaattata tgaaacaaaa ctgacagaac taaagagaga
66060
aatggccaat cccacaatct ttatctttat caggtgattt atcttggtga acattccttg
66120
tgctcttgaa aagaaagtgt attctgtagt cattgggtat aaaattctat atatgacaat
66180
gaggtgattg ataaaattat ttagattgtc tatatcctaa gttttgtaga attatttcat
66240
gaattactat gacaaggatg ttaacaacct acagctatga ttgtggaatt ggctatttct
66300
ctctttagtt ctgtcagttt tgttccatgt aatttgaaac tctgttatta aacacataca
66360
ttcatgattg ttgtatcttc ctgatgaatt ggttccgtta ttatttatgc aatgtcccta
66420
tttatctctg gtcatattct ttatcttgaa gtctttttaa ctgatatgaa tgtagccact
66480
tcatcctttt tatgcttacc atttgcatag tttatatttt tccattatct tatattcaca
66540
ctatttatcc ctttatactt aagtccatgt cttgtagaca gtatgcagtt aattgtgtct
66600
tgattatttt tactcctttc tgacaatttc tgcctttcca tataatatgc ttatcaatac
66660
agttggagtt aaatctaccg tcttgttatt tgtcacatct cccatctttt gttgttgttc
66720
ctcatttcct tgtttattac cttcttttca gttatttttt ttttgtattc cattttaatt
66780
cctcaattgg ctttatagct atatatcttt gtattatttt ttattgtttg ctctagggat
66840
agcaatatgt atacttacca cagacaattt agaaatcata ttgtaccact tcacataaaa
66900
tagaagaagc ttgcagcagt ctatgtccct ttacactccc attctttgtg ctattgtttc
66960
cgtatgtatt acatcacgta cattgtaaaa tccacaatag agtgttataa tctttttcca
67020
aatccttgtg tgaattaaaa attttatgag tagaaaaata catataacat tttattctta
67080

cctacatact taccagttct gctttctttt cattcttacc tgtttcagtc ttatctgtaa 67140
acccgttttc atttggtgtc atttccatta gcatttcagt gcagaacttc tagcaacata 67200
ttctctattt ccatgtatct taaaatatct ttattttgcc ttcgtttttg aaatatattt 67260
taattggaca tagaaatcta ggttggcagt tttctcttat actcttgggt ttcattgtct 67320
tctgatttct gttgtttatg aggaaaagtc attgattatt tgctctttct ctatacacaa 67380
tgtattattt ttctttggct gtttcaagat atttttctct ttatctgtgg ttatcaacac 67440
tttgattatg atggcctaag tggtattatt gttgtttgta tttattccac ttggtgttcc 67500
ttgagcttct aacttctgtg agcttttttt ttctcagcga atttggaaaa atttaagcca 67560
attattatat aatttttctt ctccattctt tctactctct ttggaactcc agttgtacat 67620
aggttagact gcatgacgtt gtcccataga tcactaagac tctgttcatt tttcaatttt 67680
tttctctatg ttcttcagat tggacaattt atcttgatct ctattaatgt tcacttatcc 67740
tttattatgc caccttcaat ctgatattaa ggccattcag atctagaatt tctattaggt 67800
tattatttat agtattaatt tctctgctaa gatttttttgt ctgttcattc attatgacca 67860
caatattagg ttcttaaaca tattttaata gctgctttca agtccttgtc agttaattcc 67920
atctgagtca tcttggggtt attttctatt gagtgatctt taccttatct gtcggtcaca 67980
tttttttctg tttcttcaca tgtctagtaa ttatttattg tttgctgtat attgaaatga 68040
aatattataa acagtatcaa ttacattatc ttccttttaa gggtattgag ttttgttctg 68100
gaagtagtta aattactagt agaacttttt gttcctgtca aacttgatct tattctttgt 68160
tacagtgagc ctattttagt tttaaagtta gtcctagggt acaactcttg ctctattgta 68220
tgctccttac ttctatcaca tttatttcta ttgcctgaga tagtcaatga gttctcacct 68280
gagcaggaac tgcaacattt cttgacatgg tcttacctat gtattcatca ttcatctctc 68340
aggcctgtaa gaagagatct ctgttgggtc ctgtggaatc ttgcttgcac ttggacagct 68400
cagccttcag ccaaagactt gcaggaaaac cccatagaaa catctgggcc ctctcaatat 68460
ttgatgttta ggaagctaaa cgtcaagtat agcctccttt tctagggacc ctatcttgtg 68520
aatttcactc accttaacaa ctcagaactc ttatcttctg ccttctcagg ggagctaaac 68580
tgtcactttc tgtgggctcc atcttcctgc tccacaatag gaaagtatct gcagagaaaa 68640
ggctggacaa ttgtgtagta attgcttcac gcatttccct tctctcaaag attgtaagtt 68700
tgcactgttt gctgttcaat acctgaaaat gatttctaca aattgttttt ccagttttat 68760
gattgttttc aatgggagat catttctagt accagttcct ccatcatggc cagaggtaca 68820
agttcaactt ggatcatttt aaaaatacaa actggggcat gtcacttcct gccccaaacc 68880

ccttggtagc tttccattgc tcttagaata actttgtgat ctacaacatc ttcttcaagg
68940
ccccgcatga tacaaattct ggctatttct ctagtttctt attgcaccac cttgtccctc
69000
atccaccttt tttttagtct tctctctttc tttgaacttc taccaccagg ttttttcaca
69060
cgttcttctt tccccattaa caatgatcca ccattctctt tctttatcca ctgttactca
69120
tcctcataac tgaaacatca tttcctaagg atggccattc ctggttcagt cagtctatat
69180
ttcatccccc atcacatact cttgttttac cctatatttt tccttcaaag cacttattta
69240
agttgtaatt atgtgttgtt tattttatgt ctgtctgccc tcacagaatc cacagtccag
69300
gagaacagaa atcctgcctc ttttatttat accacatcca cagtattatt agtgcctgtc
69360
acctagtagg tatgcagtat gtacctattg aataaatgaa ttgacttctg tcttttagat
69420
cgtctactca ttttatcatt gatgacaaac ataatacctt acattcgtgt agtctttttc
69480
actcctcaaa gaggattttc tgcatagctc ctctgagcct cacaaaaccc tttaaggaag
69540
attgtgaata ttatcagata aagattgtga gacacagaaa agccagatga tttggcaatg
69600
ctcatagtac cagaggcaga aatacagcta gaacagtctc ctggcctcta atcaggagtt
69660
ctttccagaa cactgcttca tcttccattc tcttgggttc tttctatcct tactttatag
69720
ggcaaaatgt gtgcaaagta taatccctct tttgcaatgt gtttttagtt tttcagattg
69780
gaatcatgta ggctttttat gcccttaata aatatcagtg agcacaaagg aagtcctgtg
69840
agggcttata atcattttgc tcccattaat tccaacactg agcagtttcc ccatttccat
69900
tcttggcctt gtgaagctct ttgctatccc tgttaaaatc taaagttgct tgaaccttct
69960
tattgcaaaa atgcatctta aacattctaa tacctctttt ttaaaaaacc aataaagact
70020
acgtcaaaaa tcagccatca atcgagaagc cctgcagtca tttgtgtgct gttgtcccta
70080
agtagaagtg aatgtgctga gctctgcatt ccccacctag ctcctctgtg atcagggtgg
70140
acattcccag gacaactggg ccgaggctgg aaacaccatc tgaatgtctg accacacaaa
70200
gttgagtggc tgatccaggt ttaaccttga cctcatcagc accaccttct aagcaacact
70260
ttggctcaga agcccagtta tttattccaa gggatgattg aatgcagtgc tagtgtttct
70320
tcagggcttt tgaactcatt tatttatcca gtcatttata aaagatgaag aggagaacaa
70380
ggtaggccaa agtggctttg tactattaaa ggctgcttga tttctaagta catgttcttt
70440
gccacctttc tgccattcca cattctagaa gccatgggta agtcagcaca gggatcttaa
70500
catgataaca ttggttttag gaggtctcgt gcataatgga ccagacttag agcacaatgc
70560
tgtaaggtag tgatttaggt gagcagcaga ttctggcttt aggagtttat tatcagatgc
70620
tttttaaacg acttgtggcc caggatccct gcacccatgg gaagcattgt agccttagaa
70680

ctctgggaat tctgaatata attcctgaat caatcgtaag gatgcatatc tgatgcttag
70740
tgcaaaccaa gaggcagaat atttgcaggc agtgtatcct tgaaaaacaa atctaggtca
70800
ttttcctgcc atgcttcaag cttacttttc catccttcct gatggtagta ctaactacat
70860
ttgtagacca tttacgtggt caacactgtg ctaagctgtt agcttcattc tctatgagac
70920
aggcactctt agcccaactt tacaattggg aaaactgaga ctcaatgaga taaagtaaat
70980
tctttacagt cattatgcta gtccatgaag gagctgcgat ttgcaactaa atctatctga
71040
ttccacagtc tttgctttta accagaggtt agcaaactac ttctgtaaag ggaagacagt
71100
agttatctta atctttgtgg gcaacatagg gtctctgtaa cgtattcttc tttctgtcac
71160
aatcttctgg aatgtaaaaa acatttaaaa tttacaaacc ttacaagaac agctcatggg
71220
ctaaatcgga cctggattta gtctgtgaat catagtttgc tgaccccgct ttttaaccag
71280
tatgtaccct ccttctcggg atgtgaaaaa ttagtgcaat tgcaatggaa aatagcaaga
71340
aaatggtaag ggcctggaag aggcagcagg attacatcag gtgctatccc tgctctggtg
71400
agatgaaact ggggatcatt gaaccacctg gcatttgtta aagagttctg ctttccctct
71460
gagattcttt caggaacctc acacctctag cagccccgag aaccgtgggc tgcaaggaaa
71520
tgcctcctca aaggagtaga aaacctgcag gatagaaatc atcacatctg tctggctttt
71580
ctcaaccttt ctcttctgca ctttcttgga tataatcaaa gcactaccag gaactccaga
71640
gtcggcacct tttcattttt gtgttttcat ttaattattt ctcagctgct aagtgtttga
71700
ctgtttaagg gactctagtg gtaaatattt gtctttagcc tggcagaagc tgtggtttcc
71760
tttgatgagc tcacacggtg tggcttttaa gatgctgctg accaggacag ctgactgtcc
71820
ccagtgggtg cagtccccag cagtgggctg gaccccttcc agaaagcgct gctgggccaa
71880
gaggcttcct ccaacttccc gctgccccca tctaaccaac acctcagtct cttctccacc
71940
tgcttccctg ccctcttcct ttccctcgca gacactttct tctgcctggc aaaaggaatc
72000
ttgtttccat ggaagcctca ttaaatctgc atcttgctca gtttgggttt gatcacggct
72060
gccagaagta tttttagccc atgcagttgc gtaatgagat agagattggg gaaaggggga
72120
ggtgactgta taggcagagg gtttttttaa aaaaaagtga gaaagagaag gaaaacctct
72180
aaagaaaaga gttttatgga attggaagaa ggatggagca cctcttttgg gagcatgagg
72240
ctggtgttct ctggttagct cttcccactg gaagcccatg gacacttgcc ataatacctg
72300
tcctggtcac atgtcagggg aacctctgat ctccctttcc atgagcttag ttggcccagc
72360
cagggtgaca cttatgctag ggagtgtgat tgatgttgct gcttacagat ttcccctccc
72420
acagacctga tggggcagcc aggatagtgg cagagaagaa gacagagcaa tagcaggaaa
72480

gagaggacaa cactaacaca ttggaggttt atgttcaaag acgggatcta gggggtcaga
72540
gaaagcacac ctaccatgta attggtgctg gaatctgatg ccaagtgcac ccttggcttc
72600
tgaggttctg agaactcttg cttgtgcttt tcagccagac tatgccctca cctgcccctg
72660
tactttaaag agctctttag gctggagtgg ttgtttgcat tggattgttg gagtgtgtgt
72720
gcatgttgtt gtgttcttgt attacaagac aaagagatta aaaaaaaacc acatgcagct
72780
gtcacagcta atgtttattg aacttttact atgccacatg gtgttttaag cattctatat
72840
gtgttaactc attttcccta attctatgga ctagacactt aaacagtctc cattgtacaa
72900
acaaggaaac tgaggcacag agaggttggg aaactcattt gaggtcctcc agctaattaa
72960
tagtggagcc aggttttgta cccagacaac ctgatttgag aatctgcagt cctagattag
73020
taacgtgttg ttggcctgtc acacatttta aatgacattc tgtacacaga accatttata
73080
gtaactttgt attgttgagc tgaaagcagt ctgcagatgt gctgctggga tttcattcat
73140
cttcaaagag gtgttttttt tttttttaa aggaaaatgc ttttctgagg gtggtatcta
73200
aattcataaa aatctttacg atcaagattt tcacaaattt cattctgact ctgttgcatt
73260
gcccttcttc ccatattccc agttagtttg tattgattgc tgcatctccc ttgagcccat
73320
ggtcccccac aacatttctt gcagaactgt gtcctgcctt cacactgtca ggcagcagga
73380
gcctctctag cggccagccc acagtcctgc agctccttcc tcaggacgtt taatttccca
73440
catttctatg cagttacctc acagaaggat ggctacgagg gcctcacttg gcttggcaag
73500
ttggtcccct tttactcac aagactctgt ttatctcttt gtttatcttt gtttatctct
73560
ttgttgacct gcccctcttc aaggcctcag ttttctctga agtttacagc ttccctcctc
73620
atcccgcaaa agaccaaagt ggaaaagatg aaaccagaat ccactgcaag ccccacctgc
73680
cacagcctct cctctaaatg cattctctgt tgtgtttagg acttgagaat gaagagggac
73740
atgaattgag gatttgttta ttattcttta caatatccct gtgagctgag tactgtaaat
73800
acccccattt gatacatgag taaactgagg tgtggagtga tagaggaatt tgctcaaggt
73860
cacataacta gtaagtgggt ggagctgtga tgtgaaactg ggcagtctga ttctgggacc
73920
tgtgctctta atcaccaatc tatattgcct cctacttgaa aacatccagg gaaaatgttg
73980
agatagatca gctgaaatct tcttgcacag taaagcaggg gccacctgtc ctggagttac
74040
attcatcttg ttcattgtca acgatttgtg ttcagtgaca ccctcttcag cccaagaact
74100
tacctgggtg ctgtgacaat tggacatgac taggaacaac cagtgacatt gtagcccatc
74160
caaacacagg gtaggaagtg gatgcttgtc actctctttt ggttataaga agcaggaacc
74220
cagtaaaggc accttttata tatctataaa gttgaatata taagatatat gggggccagg
74280

cacagtggct cacacctgta atccgaacat tttgggagcc caaagcaggt ggatcacctg
74340
aggtcaggag ttcaagacca gcctgaccaa catggtgaaa ccccatcttt actaaaaata
74400
caaaaattag ctgggcgtgg tggcacacac ctgtagtccc agctacttgg gaggctgagg
74460
caggatactt gcttgaaccc gggaggtgga ggttgcagtg agcagagatt gcgccactgc
74520
actccagcct gggtgacaga gcgagattcc acctcaacga aaaaaaaaaa gaagatatat
74580
gggtatgtgt agaactcaca gaagggcaaa caggccttaa caggtgctga aaacaggaac
74640
tgggaagttg ccagtacctt cctgtctttt cccctggaac caaacggttt cttacttgct
74700
tctctctgca cctctgtctc atttccctct ctcttcagat gatttttcat tgttgcatca
74760
cacacataga aaaatcagga tccaccctcc caagtttaca tatcgttgtt tcaggcagcc
74820
atagtatcct taaaactcca cattccaggg agaaagcttg ggtcaaggat tcagccaaag
74880
ggcagcgaaa tggagtaaag atgcaactgc caggtctatg ggcagcaagg aggccgggaa
74940
ggaagccgct gttgtggtcc aagtgacaat tcaacagctc aaagcataag taagttgtgt
75000
gcttttcaca gatggagaaa ctgaggcaca gaaggaacct ggctggggtc caggtctctg
75060
gcctttgtgt caatgctagg tcactggatg tggcgtctga tttctacagg aaatgtggtt
75120
tctctacttt gtcccagagc ccactcagag cactggctgg ccagggggtc ctagggccct
75180
cttaggatag tctcaggcca acagccccag gacagaagca accaaagtga agttatgaaa
75240
gaaagctctt tgctgatctg tcaatggcac ccttgtagag ccaatactta gaacacctgg
75300
atttgaatac tcatctccaa aacctgtgtt ctttctacca cgtgacaagc ccttgtaaac
75360
ctcacaacgt ctctatgagg tgagcgcttg cagatccaca ctttagataa gcaaatggag
75420
gctcagaggg taagcagcta gttcaaggtt atgcacctga gccaggatgt ggacacagct
75480
ctgtgtctga ttcctaaggg cctgtgcttt agccactttg caatactgct gctgtctgct
75540
tcatttcctc atctgtcaga tgggaacgat aatactcaac tcacatggat actgtatgag
75600
gaaaaacaga taaaagaaga gaaagtgctt tgaaaacata agcagccctg gcagatggga
75660
attatttttg ctgctgacac acatcctcag ccttgagggc tctgctgagc catacccagc
75720
tcagagctct ggaggcacct cctccccatc aacagcaggg gggacattct gtcttcatcc
75780
tgagcaggct gacaaactga accccactcc tccctcaatg tccccatgct gggaaggagt
75840
atagctcatg ctgtgttctg tcttgttgct gagagaatgc agaacccaga atttgggtct
75900
cagcaggttg gggagaaaag gaaatgtatt tcttccccca agatttcttt ttgaaatatt
75960
ttcatttgtg gaatcagatt gtgcatgcaa gtttcttcca gaaatgtaag acgtcgtaat
76020
gatgggaact gttggtttta taattgaagg atgggaaagg aaactgatat ttatggagca
76080

cctgttctat accaggcagc tacccaacca tcagccattg ttgcaatgtt atgcaagctt
76140
tattatccac atttcacagt ctgagtctga ctcagcaatg ttgtgttcta tgtgctagtt
76200
cccacaggta ggtggctgca gcgctgggat ttgaacccat ctccaaagcc tccatctttc
76260
taccactgcc tcccattggt ggggaggcca tggactggct gtcagagatg tcctttccag
76320
tctagcagac taggaagctg ctggaagcta cttatgcaaa ggtcagcaag gaaggaaaca
76380
gagtcagaac tagatggggc tcccctggcc acttttccat gctggcccac atgtccggct
76440
agcagtcaac attgggtctt atgcagagcc acctgtgttc aatggaaaca tcctggacac
76500
tgcacaaact agtgggagcc tgtgagggaa cagcctgtcg ggttcattga ggttcagccc
76560
aactcatgag ctagggcagg taccagaggg tgtgttccac ccaaatgggg caggtaggca
76620
ggggacacag gctccatttt catgaccaaa gactgagcag agaggctctc tgagcagtgg
76680
cagaatggga agtgtcaaga agctttgttt gacaattgag tcaagaggac agaaaagaca
76740
gaaagcagac atcagagttg ggaaggctca ccccagctcc ttgacaaagg tgcatgaggc
76800
cagttcttga agcagtgacc ctgccttatg tcatgtgttt atcaaagccg gcccatcagc
76860
cctgaagtgg cctctgtgtt tagaagaggg cctgacatga ttctctgaga aaggatttga
76920
caacaacaaa gtgttgccgt atgtgttgtc tcatcccctc aatagtcctg tgaggtatgt
76980
gagacaggtg ttactctctc cacttggcaa atagggaaaa gagggcccag agaagtgaag
77040
ctgctttccc aggaccacac agctggtaaa cagtgtccat ctcagctgtt ctgtctccca
77100
caccaaatac cctgtgcacc acgcaaacac aaagacaact ggacaaccaa gtcatctaat
77160
gagtatgcat gctatggtct ctctcatttt gtctttcagg gctataccct aggagagcta
77220
atcattcttg gttagataag aaatagccaa cacttctgca gcatggtagg ccaaatacca
77280
ccagaataaa ctcagaccca aagagatgct cagaatgtgt ggagttaata cttcactata
77340
cagctctaag gtataagcct tgtccatctg tcacattatg acatgtgctt gctcccacct
77400
caattcctga ttccacatta caacaaatac aatttcaggc tttgaactaa caatgccaat
77460
gtttctgaag cccatattaa atgccaaaat ctgagtcagc tactggaggt agagacatga
77520
ataagatggt ccatattatt ttagaggatt ctttggttgc aaagggcaga cacccagctt
77580
gaattcactt tggagaaatt gggatttttt tggcttgcat aagcaaagca tgagaaagaa
77640
agttccaggg atgatgaaaa ccaggaatgc aaatgtctcc agaattcttt cttttttctt
77700
ttaggccatc ttttttctct caaactggtt ccctccactg ggctggagac gttactacca
77760
gcagcactca gacccacatc ttcagtttaa atgttggaaa tggactgtca gagaacattt
77820
aggccattca ttctgtggga gagataggct atgtaaaaag atagccactc ccatgtgaac
77880

aatgtggtta ggattagagg catgaatata ccccaaacca ggggtgtggg aaggaggttg
77940
acactctagg tgataatacc cagaccttaa ggagctttct gtctagaggg aggtatggac
78000
atggacaagt aatcaacagc tacaaagcag agctgccagc tctgcaacac aagagccctg
78060
agaggcatga caggggcagg gtggggatcc atgtgggtct ggattgaagt gaggaggggc
78120
atcaggaaag cattccagga gagctgaggg acacttgagc acaccctcaa agaatgactg
78180
ggggtcatga ggtatacaag ggaggaagtg cacccgagac agaaacaatc acataagcaa
78240
aaatgcagaa gaatatgagg atcggggaag ggcaagtagc tcagtagtgt tggaggccaa
78300
gggacacgaa ggaaggtgat aaagccctga tgttaaggat agaaaaatca aagtcctttg
78360
aaaatcatgt ggagttagga tctcaagaac cctacaagga tttctttaga atagaatcaa
78420
agaaaaacaa agtttacagt ctgtgagggt tgcataggaa gtaacgtggt gagaaatgtt
78480
ggcttgagaa ccacatatcc ataacacaat ggtgttttag aggatttggg ggaagggaga
78540
gaaaatctca aattgtctca gtaactaatg agctttcatg tacatttaaa atagtaataa
78600
atgcaattgt gaggatgatg gtgagatgag caaaataatc cagtttgtaa ttgtagttat
78660
caggctggca tatcctgcag gtcacacttc taaacatgac ttcgaaaaat caaagatcag
78720
ctaagtttga agtaagtatt gaaagaggga gattatgttg cctcaagtta aaatagaacg
78780
taaaagatgg tgattcaaat gatcaaaagc accaagcttc cctgttagga ttcaagggag
78840
gggtgcgtgg ctccgacacc agatatctgc aaagcaatat gaaatgagat caatagtaga
78900
cattgaaaga ttgaaactga tataggatat tcaagtacca gcttcaagaa aatgaaatga
78960
gacctaataa aagagagtag gagtcaaggg ggtatacgat attaaagaaa gtgaagagcc
79020
agggtttgta ggaaggaagg gagaagaggc aaagagagca gctcttttaa cacaggagct
79080
tcctcctttc ccattctccc tcctgctaaa agccgagttt gttttagctg aaatgattgt
79140
aagacaaatt tttattatta aaaaaggagc tattttgtgt tggtttccat tataaaatca
79200
gagctctgct gccataaaat taaatcccat aataaaatga gtagaaaacg tgatgtcctg
79260
cagaaaggaa gatggcagcc cactcagtgc catgctgggc ttgactatat acaagccgtg
79320
catctcctgc tgcgagttgt agctgctgcc cagcagtgca cattatcgtt gcagctgttt
79380
tcctcacatt ctgaggttta tgaaatccct catccatcaa taattgatct ttagctctta
79440
gtccaggggt tgtcaactgg cactccatgg acctttagag gattgatggc taggttttca
79500
aagatctttg aaccccctga aattatatac aaaatactgt gtgtgagtat gtgcattttt
79560
ctggtaagaa gcacctgaat tatcgaagca gtttgtgatc ccccaaaaag ctaagaacta
79620
cttcctagag caaagggaga ttttgctaca cttagagatt tacacatttg accagggcag
79680

ctcacacaag tgggatgcgg tttcacattt catggcagat ctgcttccag ctatacaaat
79740
tcatcaagga aatattgtaa tacttctata tgaatcagga attcactata tttaacttat
79800
ttggaataag aaccactata tatatacaag tttttccaaa agactgaagg ttcttcctgt
79860
ggcaggaagg aatatgatta gattcatgaa gcgcctttat gtttatattt caactctgaa
79920
agataattgt gactttacta aatcaaacct gtataccacg attaggaaaa tgtggactga
79980
tttggggttc taggggtaaa atgtgacccc tgtgaagtac caatgcaccg ttcttttatc
80040
tgtgaacggg cactgagctt ctgaaattaa ttagtaggca ggaggacatg cgcatatgac
80100
gtgatagttt aagtactgat aattattcac ttggaaggga agagaataaa attcagaaca
80160
cagtattcct taatgggaaa tcaacttaga ggaggtagga gggagatcaa gcaagaatat
80220
ttctggtaaa acatgcataa atcaatggtc agccaatgtg ttgatcaaag aaattatctt
80280
tcggggaaaa cagtagaagg caattgaaaa acaagcatca ggctgcataa aaacagcaaa
80340
caaaagtcac aatggcttga ttgtgtgatg aggtaattaa tggctgcagt tagcaaaata
80400
tgttcaaaaa aaagacagaa agggtagtta caggagaaaa acatccccgc agatcttcaa
80460
aatcagaaac aatgaaaata attatttcaa aaattaagaa aaaaactctc taatttatac
80520
ctgaattacc tggataattg gtaaaatttc ctgcatatac aaatcttggt cctctgctcc
80580
tctctctata aataaataga aatgtatgaa tcaatagtca gccaatgtgt tgatcaaaga
80640
aattatcttt tggggggaaaa ttggtagaag ccaattaaaa aacaagcatc atattgcatg
80700
aaaacagcaa acggaagtca caatggctcg acggtgtaat gaagccacac aatatgtatt
80760
aaacacatca tctacacaga tggattcaaa gataccttct ttgtgtctaa gtcccaaatc
80820
tgtgtttcct ggctctgttc cctcatatct agtcattctc caagtcagca tgcccaactt
80880
gaaagtgtca ttttcaaaac ctgcttcttc tcttctggaa gttcttcctc tgcccattgc
80940
tccacaatcc ccacctcttt cacccagtag caaaccttaa atttatcttt tactttgtct
81000
tacttcccct tcttatattc aaaatgtttc tcacttgcat ctcttttcat tcatttcata
81060
agcatttatg agctcctgtt atggtttgga aactgttctt catgctggag gtggtcttat
81120
aaacaagtaa tttcaattga gtatttagta tgttaagtgc catcccaaag gcaaacacca
81180
gctgtgggag gctccccaaa tcagtctaag gaagttggga aaagcatctc agagaagatg
81240
gtgtctgaga tggggaggat gtgtggaact gggcaaggaa gagaacaagt aacaacattc
81300
tagaaaaagg cctctttcag catgctaaga agtttggagg acagaggagt taccattcaa
81360
aatttggagg gaaggaagag catactgagg tttgccactt gaacagataa tttcagctgt
81420
gttgggtgag tgaagttgag tgggtacaaa tcaggtcagg aatataagtt aggagactgt
81480

tactagaatc caggccagag gtgatggtgg ccaatatatg agagttttag cagggaatga 81540
aaaaaagaaa atgtgttcat gaggtagaag taggtaaaaa caacaggatc tggttcctga 81600
ttggaaatgg gggtagcctg gagaggaagc cagaatgcag gcaagaatgc atagtggtac 81660
catccactga catagggatt aaaggaggag aagaagcttt ggtaaagaaa ataagaagtt 81720
cagctatgga atgtttgaat ttgatttctc tgatgaggag tagttctagg tgatgataat 81780
gctcagggtg tagacttgag agtggatggg taaagtaaag gttgaggcta ttaaaaggga 81840
aaaggtcaag gaactgaggg ccaaggattt ataataagtt atcttgggcc actaaagcca 81900
cgcaggatgc tggcaggaaa cctatgagcc aggtcttcaa tgttgagtcc agtgactcag 81960
gtgtcagaag cagcaggaga agcattgata gcctgatggg gaaggagccg ttacctgaga 82020
gtagcagaga gagttatcct agctgacaca gctctcaggg atttgcttct aaagcaatcc 82080
ttaggaaaga aagagcagta tccacaggag actggtgggc actggcttcc ccagaaaacc 82140
tacctagatg aattctattc tcaagggact cctatttaga taaggggctt tgttagttct 82200
cagagcaaca ccaaacagat gtatatctca ttacttgccc ccacaacctt tctgctctgg 82260
ccacatgggc ctacccactg tctgctaaat gcacttcata ttttcttgtt tcagtgcctc 82320
agtattcata atcttctttt cctaatctct gcccctcact tacctgaatc ttttgtattc 82380
tcaatgacct gctccatccc agccctttca agaaccttta atacctacca agtgaatact 82440
ctctccattg attacacact tcctgtagca cctgttctat aattatgaaa tattacctat 82500
tgtacacata tatttcaatc tcttggtgga cagagaatcc aatttatgcc ttgtcaattt 82560
gtagcacatt tccttgcata tgtagatgca ccatgaatat ttagagaact tgttagttaa 82620
tttcctgttt aacatgggct gcaaagttct ggtccatgca cgtcttttat aaaatagaaa 82680
tgacggatgg tgcatggagc ttaaattcca tgaagcagaa acatatgaga gatggagctg 82740
aatttgtttg cctgtacagc tcttacagca attgcttcca atttgtttga tttacctaag 82800
agctaaaatt gtaaatggca gctcaaatga tttttctgta cattcagaaa atgagtttga 82860
atatttgttg gagagtaact gcttaagaca tgaaaaaggg ggagattata gcttttaact 82920
ctttttatg gcagagcatt aaggaaaaaa aagtgcagat aaatgagatc aaatggcaag 82980
tgtctgaacc tgctggacac aagtcccggt agccattgat agacagtgtt tatatgactt 83040
ctgggccatc aatagataga taaggtacat cagcggccaa tgttccagga agtttgagaa 83100
gataaatgga agttgcacag cagcctaaaa gcttccttag gagggctgtg ctcctccaga 83160
gcgccatctg cctgtgtctt cctgttcttc ttcttcacat taaatgcttt tccttttctc 83220
atttttatga tggttatcct aaagatatgc tagcctggac tttgacaagg acatctggag 83280

ataagaaaga ttctgaatta tttttccctt tgggcaattg tagcaatttt aaaactatgt
83340
tagatggcta gagattcttg agaatatttc ttttcttgga aaatcataag gctttggata
83400
gtggtaccta tagaagctga catcagcagc agcctgcctc cagtcgatca gggcctttgg
83460
aacttcacgg ggctcctcta ctgacagccc catcggtttc cctccagcac acgtaactca
83520
gcattgactc tgggtagtag agggtggttt atggaatctg attcatctca gaaagaggtg
83580
gatgcaaaca cattcccaga gcagaaggct tggcatgtct ggtcttaggc agagggaact
83640
ggagatactt gtcctattgt tcttgagatt ccagcaaaaa tagcccatta cagaggaaga
83700
agatatcagg tcaaatgaag gctttggtgc tacaacattg tcttagaaaa aaaaagaaag
83760
aaattggcca agtgcagtgg ctcagcactt tgggaggctg aggggggcag accacttgag
83820
atcaggagtt cgagaccagc ctggccaaca tggcgaaact ccgtctctac caaaaagtat
83880
taaaaaatag ccgagtgtgg tggcgggctc ctgtaatccc agctactcgg gaggctgagg
83940
ccggagaatc acttgaacct gggaggcgga ggttgcagtg agccaagatc gtgccattgc
84000
actccagcct gggcaacaga gtgagactcc atctcaaaaa aaaaaaaaaa gaaaaaagaa
84060
aaagaaaaaa gaaaagaaag aaattaaatt aaaaaaattg ttttttaaac aaaggaaggc
84120
tttgggcttg gagtccaact aagctaggct ggaatcccgg tttcatctcg cttctctgtg
84180
caactttgga ttttactgaa tctctcttat tctcaattcc ctcctctgta aaatgaagat
84240
aatgctagta cctgtctcat caagttgaag gagacttaaa tgagatgtgt tgaaagcatt
84300
tagcatagta tgtggcacat aaagaacact caataaatgc tggctataaa gaagccagag
84360
agagactcgg aggtgatgag agaggccaca attccctcca tttcattgaa aagcaatttt
84420
tattatctca tttgaaaggc agtatagtat agtggttaag gacatgcact atggagctag
84480
acctcctcag ttcactttct gtctctatca tttattagct gtgacttaac cttcttgtgc
84540
ctcagttttt atcattttg agagaggagt aataatagtt cctactctgg tgtgttgtgg
84600
agatttgatg agttaataca tataaagcac acatagtagt gcctggagca tattaaatga
84660
catgtaagta ttagctgtta ttttattaaa caacatgtgg cataggacat attggaactt
84720
tgaagtcttt gaggctcttc ccagtttcat aaatcagaga ctacagtata aatatctgct
84780
tacatgtctg ctttccccat tggactgcga aatcttgaaa ctgttttatt catctctgca
84840
tagcgttggc atcgtattat gatacctgac atttaccagg tgccaaatgg gactgggcat
84900
gttgtaggga ttcagtcaat gtgggtcatt gcaggcgggg aggtgggtcg ggttaaaggt
84960
aagagaaggg ccttggggca tcacattaag tagttaccag attgaactgc aaacattgct
85020
atccaggaga aatcaggtca atatttcacc ttcatggcaa taccagtaca gtccaaggag
85080

aatgcataga aggaaagaaa tcataatctg attgtatgtg tttttttagt agtaaataat
85140
aataattatt actattccta tacaattttg tgtgttggtg tgttttgttt tgttgtgcat
85200
gaaaaatggg gtgctaatct attccccttc ccaacaccag tgctcagaag aaatttccac
85260
agatagagaa gctataggtt atgaatttgg ccttgatgga ttctgggtca ctatttctca
85320
atgtttgtcc atgtcatgtg aagctcttaa gataaagaac aatgtcttac tcgtcttttt
85380
aacttcttta ccccctaatg cctatcacat actttgccca tggaaactca atagacattt
85440
gtaaatggaa tttaatttct gaggtccagt aaagcctttt tccatccttc ccctactaca
85500
cagtttgtct aaccatgtct tcccttccat catccacctt ataaacgtta ttactcattc
85560
ttccatcaca ttcttgacac ctcccatgtc caatgtcaaa caagtaccat ttgggaaaca
85620
gaattctagg aatctggaga cctagagctc ttcagaccct gaaatccagt tttctgagct
85680
gagacagttt cttaatttct cactccaact ccgtttctcc tctttctcaa tggatatttt
85740
ccaagtctcc attaggcata tagcaattcc agaaaacatt caattttccc ttctcttaat
85800
gccatgctcc aaaacaccac attccctcta gacattgagc attggagaga gatggaaaag
85860
tactttgaaa atgtgtgcat gtgagaaaaa tgctaagtgt tctgtctggt cacttcaatg
85920
acaagtttgc tactttagaa acttgactaa acagagtgtg aggaaaaaca tgaaaagaaa
85980
aaaatgtgtt cagcttggct gaataatgac cagcagggtg aaaagataag ataaccaccc
86040
gctcacagga tttctatcct caagccctag aaggttgaca acagcagaca ctgaaactac
86100
tcttaatgga gggtgtgcta aagaagcaac attatagccg cttttaggaa agcaaatagg
86160
aaagttggtg aaatagagaa gatgcctaag catgtgagat accacctcca tcttggaaaa
86220
taaccaaggt gatacaatgt tatgcaggac cccttaatta aaacagattt agtgattaat
86280
atcaggagca ttgtcaagaa tcacaacaac agcaattagt tactattgag caatttctgc
86340
taagtaattt gcaggagggc atctcactta attatcacat ccttttatag atgagaatat
86400
agaggcttaa aaaggtgctt ttcccaatgt tattcagcta taagtggtca gtcatgactc
86460
aaacataggt caacctgaca acaagatctt cactcttaac ttctcttctg tgttgtaata
86520
cccttgatcc atggaaatgg accatcttca tatactgctt ttttgcctct ggaatgtcca
86580
ggtatggatt gggtaatgct caaagacaga gaggaataga gtattaaaaa gatccctggc
86640
ctcattttct gaagacatga gcctaagctg agctgtacca tttaccatct atgtgaactt
86700
gggcagattt tttgacactg ctgggtctca attcctgtaa ctgtcaagtg gaagtgagcc
86760
taactgcata gacttcactg ggctgttaag agaataaaat gaaataactg taaacagaag
86820
tgcctagtgc acatgcaaag gattattggg gctttctacc cttcagggat tagaagttga
86880

tagtaggcaa caagttataa gaaatacagt caattgtctg ctgaccaggg ctagagttaa
86940
ttgtctctgg aaaaaaggac ttgcctctct ttctcttctt cctccaaaac ttaagacgtt
87000
tgcagctgaa tccccaacag gattttgttt tcctttggga gagaggaaac agaccaatat
87060
acccccaaaa ctaaccccat aatttcattt cagcagtaaa gtgaggtcct tgataactgc
87120
cctgcccaac ctgcagggtg gttgggaaac tctgaatggt catgcatggg gaagcattgt
87180
gtccactgta aagagctctc cggagatgat aaatctcatc agaaggcttc atgcttgagg
87240
catggattct tggaaaaaca atcactctac gtatgtggtc agaatctaaa ggagatgctg
87300
gggagaggag ctaggtcagt ctccaaagtg gaacagtaga aactaatcat gtggagccta
87360
aacttatgaa ggtttttaaa atcagaattg gccaccttcc tttggaccat gagctcagat
87420
tgtgaggtgt gactaggtca cgtctccttc ctgcccctgt ttccctcctc tccctacctg
87480
tccctccttg accccaggaa aaattgccgg gatatgaaag ttaattatga cccaagggaa
87540
ttggtacaga tggggaagaa agaaatgcat tcaagagcat ttccatcagt attgaaatta
87600
cacagaaggc tggtgaattt gggctatcca ttcttgcctc cctctgtgcc cataattcct
87660
tggcctcctt caatttcatt ttccctttgg ttcagaggaa tgcttgatgg cttaagctag
87720
cctcagttgg ccaagcattg gagaaacaga gaggtgtatg acacagctac actcccatgg
87780
ggcttacagg gcaaggtgag agaagacaga agttgtatgt gctgggtgcc acgtggtagc
87840
tacaaactag aaatgagacc aggttcggaa gaggaagagg gcttgcagac ctgagtcatg
87900
gggacagttt cttcaggaaa tgggatctca gctctgcctt gtatgcaggg cttacataat
87960
aaatatgttt cattgttgtt gttgttattg ttgatttaat aagattttgt tttaagaaga
88020
ttttgtaaaa acaactgaac aaatgcaatc tcctgccaga gcaggcagca gcaaaggaga
88080
ttaggaatat aaccccccttg gagacgttcc ttcacctacc tggtgctgga ttacctaaaa
88140
gcttcagcta agtagggtca cccccccaag aaattatttt aaaaaaattg aaatctgata
88200
tttttagaaa atcttatcaa ggatatttaa ttggactatt tacacctatt tagggtcagt
88260
cggttttgga caagtatgca ggggtcttgg aatcagacca ctggggtcaa atcctagttc
88320
tgtcacttcc tagctgggtg accttggaca aagttacctg acttctaata gcttcagatt
88380
cctcatgggc aaaatagaaa tgctactagt acttaatagt gctctgagaa ggattcaatg
88440
agaaggatta aatgtatgta aagcacagtg tttgcccata ggaagctgtt atttataagg
88500
gaggggagca tcctaaggtc ctccgaattt aggagaacta aaaatcttac actgacttct
88560
cccttcaaca gcacccttcag aatctccttc atttttcata ctgttctttc aaccctttga
88620
tgaatgagaa attaggcatt ctttccctgc agattttccc aaaccttctg ctttggccaa
88680

taaacatatt tttagtccca atcttgcatg ctcctttggg acttttcatc tgataaacat
88740
cccctcctgt gctcttgaat ccaataccct tcttccctgc cctccaccca gagtctcctt
88800
gtatctgctg ttaggcacaa tgatgacccc accaaggtca gacaatggct gtggcctcac
88860
ctggaccttg atgacccaca tagcctagag cccagagatc agccactgat ggaggcccag
88920
agggcagttg gaaaacttca caagacaatc cagcctgatt gttttgacat gcctgacttc
88980
aggctgctaa aaatgagctc gaggaatcag ataggaaaaa gagataggtg atgcaatttt
89040
attccatctc ccaattttct gagtcaagag ttgtttgttt aactccagtt aaattagtat
89100
ttatccaaat ttcctgggtg cttgtccaaa gaaaagtacc ccagatctac aaattagaat
89160
ctgggactgg gacttaggaa ttggcacttt tacaattata ccagatgttt ctaatatgag
89220
tacttcaacc actaccctta tagaagtgct gcctaggacc ctctcttctg gcaggtgaag
89280
tggaaggagg ttttgtcgaa gggagattct ccacttcaac ttgagtgtct tggcttgtat
89340
ccgctttgtt tggttctatt tcaccaaagg ctttcatctt cacataaatt ttcttcagct
89400
ttaaataatt agttttggta accattggta tactggaaag aacattagat ttggagtcca
89460
ggtggcttga gttcaattct ctgctctgcc atttaccagc tgtgtgacat tgggcaagtt
89520
gccaacctat ctatgtcatt tcctcatgta aagataatcc cacttcacca ggccactttt
89580
gaggacccag tgaaatgatg tgtaaccatt ttaggaacac tggatcattc tacagtgcaa
89640
ttttttacat cagcttggag cctaccatgt aggcattcaa atccactgag tgtatggagc
89700
tccgtgcaca aataaaagga cttctctttt ctgcccgtgt acaactttgg tttccttaat
89760
caatagaatc catgacaatc ctgggccatg gtataaagat gggactttct tcctgtgaag
89820
gagtctggtc tgaacatctt ccaaactcca acataactga tgtcatttct ccacccaacc
89880
ccatttgctg tctcctgact caattgctag agaagccact taaggaaggt tcctggagtt
89940
aaggctgtgt ctgggccagt gtagcgagca gttttcaaca gtcagtcctc tttatcttct
90000
cttttcctgc gagcctttac taagcactgc ctcctcctgt ctccttactg catctcctga
90060
tggaatgcac aggtaaatct ccttggagag taccagccag gaacagtcca cagccaaggc
90120
caccgatcct caccgctgag ctccatcttt cctttcaagc tgtccttccc ctcccctccc
90180
caccatcacc atagcaacac agtggtataa aaaaatgaaa gcgctaaggc atctaaatat
90240
agtctgagta tcaactcttc cagcatggag ccgaaaacct agggaatgac agctagaggc
90300
atccagacga taactggcag ccaggagggt ggataagtca aaggaagggg tcaaggaaag
90360
aggggaagga aagggaacca tcacttgctg agcctgctgc ctgtgctttc tcatgtcacc
90420
cgcacgacaa cccaatgtga atgttatcat ctccaggtaa ctgctgaaga aacggaagct
90480

caaagaggta agagatttgg ccaaggtcac acagctataa gcagtagaac taagatttta
90540
actcaagttt ctatggcccc agaatttatg tgtttctctc tccataccac agggacaggt
90600
gcaagtgaga gattttgctg gaagcactgg gctttttgag caggccatat aaaaattctg
90660
agcccagagc tcaactaaat tattggaaga gactgggcca aatataaggc ttctatctaa
90720
gcagcacctg tgtttctcaa ggactgagga aaatgaaggg ggagggttgg caaggctgca
90780
tttcccaggg tgcgtgatta tatggcatgg gggtgggggc cattatgatg cccggacatg
90840
gaacttacac cagtgcagaa agggtgtgat tagaagccct aagccagaga atgttcagtg
90900
tgataaatgc cattattttt tccctcattc attcaataga tttttttttt agatggagtc
90960
tcactctgtc gcccaggctg gagtgcagtg gcaccatctc agctcacggt aacctctgcc
91020
tcctgggttc aagcaattct tgtggtccag cttcctgagt agctgggatt acagatgtgc
91080
accaccacgc ctggctgatt tttttttttt tttttttttt tgtatttttt agtagagaca
91140
gggtttcacc atgttggcca ggctggtctc gaactcctga ccccaagtga tccacccacc
91200
tccacatccc aaagtgctgg ggttacaggt gtgagctacc gtgcctagcc tcattcaaca
91260
gatattttta ttaagcatct gatgtgtgct taactctgga aatatagggg tgattagaac
91320
aaatgcagct cctgcccttg tagagcttat taggatagtg gagaagacaa ataaggaaac
91380
aattatacaa ttgattgatt ctttacaact gtaacatgta ctataagtac ataacagaag
91440
aatatcactt gcctgatgac ttcagtgaaa gggaaataca gaagttctta caaatcaaag
91500
caatcccctg ggccaattgt aaaggtgatg cccactttca aggtggacag agactgtgct
91560
agaagcttag cctcaaccat gggtttatat gattggtaga ccctgcagat ccattcccaa
91620
tggtgtatct tcatactaat catgaaatcc atctaatagc catacaagtg aggttttaaa
91680
acccaacaaa ctagactcaa atgaaatctg atgagggaat ttatgatttg ttcttcctac
91740
agcctttggt atcactgaca taaaactgaa tgtatgtgct gagggtgctt gtgtcttggt
91800
gatagacaag gtaggtggtc cagcccatgg tactggcagc ttaaagtcag ccagccatca
91860
gtgggaagtg cctgtgaatt atgcaggagt gggaggggag ggagtaggca gtaaagtaat
91920
gcatttctgt ggatccaaag ctttccaaac tacctgcaag tcagcaaata tgggggatgt
91980
tgtatgacta agtgagaatc agataatata atgtgtatgg agctctttag ttcttcagaa
92040
aaaaatgctg tctaaacaaa tagtgctgat atcaaagata atgatacagt accctaattt
92100
taatgctctg ctacctacct gccagctgtt tcccagggat gtggtaaaga tgaatgggca
92160
agatctggga aagtgttttg aaatccttga ttaaaggccc tccaggcaga tgtagaattt
92220
taaatgtgtt atattactgc cactattgtt atgctttctt ttatcacccc agaatttcac
92280

catctcctgt ttcaggtgaa cgagtctgcc tgactcttac ctgccctgaa tggcattgga
92340
aaggtagcag ccctgagatg tgccatataa acaaacatgt ttttaaccaa gggatcagga
92400
ggccttcctg gctggctcct gtcagctggt catcacctct ctataactct aggctttccc
92460
aagcttattt tatttccatc aataggacag gaatatgtaa atgtcctgct tgaaatgagt
92520
attggctaca agccatctgc ctctgaacag aggtgaaaag tggaaatcgg aggaagggca
92580
gatgtctttt gcaagggaaa cagactgttt tctgccactg cactctgccc aggcaaaaga
92640
gtaaaggaac agcactcagg agaattcact gaagcgaggg cagggtgcaa aaggaacttg
92700
agaaattggt actgggaccc aaaatcagat tctggcattt ctgggaaaag aaatgggcat
92760
gggtgggggt tttatctgtc aataaaagca tccagaatgg ggctagaagg aagtaaattc
92820
agttgccacc tctgcctact ggacagccac ggagaacttc tccttatcca aggtcgagga
92880
gccctccgga gtacatactg ataccattgg ttctcccaca catacccccа tggagataaa
92940
aacaggaccc tggaagccct gtccgtgttt aaccaatggg attgaaacat ggaaatgaac
93000
tgccccacaa tccaccctgt gagagaccaa agagcagtgt tggattaaca gggaatgtta
93060
ccctgaaaag gcattcagct tccactgggg cagcaggtac agtgcaaaga tgatcccact
93120
taaattccta agacaggaaa taaggaaaga tgttgtggaa actcaagacc tctcaaagca
93180
tactcctttg tagttcttcc gcagaccaga ccacggaatt cagaaaacac cctacctggt
93240
tccaaaccag cacctgccaa acttctcacc ctcttctgac cctgtcctgg gagttaagaa
93300
aaaaaaaatc actttattgg ttgctccagt tataacttaa acagacagac catcatcaaa
93360
ttaagtgaca tgtacgactg cttattgtat gccagttact gtgctgtggg gttttggttc
93420
cattatctca tttaatcctc tcaaaaaccc tgttaggtag gttttattat tgcactcatc
93480
ttagattaag gaaactgagg ctcatagaga ttcggtaatt tgtcaaaagc cctaaaacat
93540
aattactgcc tccagatgtc tctgattcta aggcccaggc tcttaatcag taaatgatca
93600
aatgaataat gattttcatg gcatctgtca tcggaaagaa caatggagaa tatgcttaac
93660
caaagtcata accaaataaa tgaacttgac agcagagccg tgattctagc caagatgact
93720
attttcatgc atgttttgaa ggccaggaaa aggaggttag acttgtttgg gaagggaaac
93780
aggagctatc aaggtgaact tttcctaaga gtagcccaat aatagtgctc gggagggagt
93840
aatgtgtgca agaatagagt cagggagacc agccaagtgt gtgcctcagc atccctagca
93900
caaatcacac actaagcatt aagattgtct ctgcagtgag aaaggcctgg gaccaaattt
93960
gggctccacc acttactggt attcattaat cattcatgca ttcattcaac aaatatatat
94020
tgcgtgtggt ctatgtgcca gagactgtgc tgggtgctgg caaagaacac agacaaggtt
94080

cctgctctca tggagctttt attctgatga aggaaacaga ccacttacag ataaataaat
94140
aaacaagata aagggaaaca gatatgatgg agagtagctg gagggccaag cagaccgggc
94200
agacaaggtg gtggcatgta agctaagaca tttaaaaaga acctggtcat gagactatct
94260
ggagaaggaa agctccaggc agaggaagca ggtagtgcag aggccctgag gcaggaatga
94320
ggacaagata tttgagaaaa cagaacaaag gcaggcatga ccaggccgag tgggtggtgg
94380
aaaagtagta gaaggtgagt gggggagtgg gggcatcaag gtcaggcttt gcaggcttga
94440
tcagcgttct cactgtggtt ctggagccag cagcatcaat gttacctggg aacttgttag
94500
gaatgcaaat tctcaggccc cacccagacc tgctgagtca caaactctgg gatggggcac
94560
ctcattgtgt tttatcgagc cctccagatg attccgagta tgctaaagtt tcagaattcc
94620
taggttggat tatgcagttc aattttaatt ttaaatgcaa tgggaaccta tgaaagattt
94680
aagtagggga gcagcatgtt ataattttct ttaaaaaatt gtttttaagc actcctgctg
94740
aggagagaat ggaccataac aggctaagag aaatggaagc agggagataa attaggtggt
94800
tattgcaaga ggccaggtaa gaagagaaag tggtttaagt agggtggtgt ggcagagaag
94860
acggttccaa gcagaggggg accacgctga caaataagcg cgggccactc acgcaagccc
94920
aacaaggcag aaggcagaag gcaaaagtga aggccagaga aaactggaca ccacctttcc
94980
agagcacagt tcaaaggcaa tgtcctcaaa gaagacactc caccctcctc ccatttcctc
95040
cctattgcct aaaaataaga aggatacgcg gcctatggca aaccttgggc aggcacgtgg
95100
gagctgagct cttgcaaagg gcagatagtt cctctggtga gagagaaaag gaagggccag
95160
tgaggagtga aggaagagac gaacagagag cccgaaaggc tgagaacgtt gtctggcttc
95220
ctgaaaggct taaggggtta gctctggagg gtgaactaaa agccctagtt atattaaaca
95280
cacacgcaca cacgcacgca cacacatgcg cgcacacaca cacacacata cacacagttg
95340
aaggagacct gcagtttcca aaaacaagag ttgtattttt tttgttcata tcatgaccca
95400
taacaatctc aaaagagaaa caatctcttg tcttccttgt ttaggcttag gagaacctgt
95460
agtaagtaag cagcagcagc ggaactcaaa ctcgactctt cctactgtca ttctctctat
95520
tacaccacaa ggcatcagag gaccactaga gtcgcctccc tagggttagg gttagggcaa
95580
ggtaaatgaa gtgagtcagc aagggcagga taggaacctg tctttattaa cattttgata
95640
ttttgtttat catggatttg ttgcattaat tgcaactttt aaaaatcatt gcattaaaat
95700
attattgatc ttgattactg agtttttagg tgtaccctta aatgttgcac ctctgactta
95760
ctagtctcac cctgatccct gtcctggatc tatgcctgtc tgttctatat cagcctcttg
95820
ctttgaccat aagaataact tcagaccttt aagcatagag gaaataggat ttctgtctcc
95880

cttccccacc tttgtgataa tctcagcttc tgcttttaaa gtctatctcc caagtagttt 95940
gcctactatg ttcctcccaa ggtcactagg ttctgtgaaa ctagcagcag gctagattgt 96000
cacattagca caaaggatcc actattcctg cagccgagct gggacaagca cttaggccca 96060
ctgactccaa cccttcaata gcctgggacc tacgttgtct ccaggtggta taaaacaaga 96120
atttcccctt tgactgggag aaaaagggaa gaactctaaa ttggaaaaca ggtcatctcg 96180
aattctcaca ggtggaaatt tctgacaacc cctttgggac ccacaattca acacacccca 96240
aatggggaca gtagctaaca tgcaacctgt aggctgttct gtcatccagt gccactgtgc 96300
tgcacaccac caggggcag cattctcatt ggcttctatg tgcctggagc ccagtgcagt 96360
tgtgcaacac tgcagctttg ctttagtgta gtccctgatg ggttcagtca agaaaatgtc 96420
tatagaatca gctaatctcc catgcagtta agtctctaat tgaaatattt tctctgctca 96480
gcccagggac agcaatcttt cctggatttg ctatttacaa ggatctctag aaattatcca 96540
ccagaaatat gggctttctc agagcttgag tggacaggga attaaggtgg aaggcagggc 96600
gttttgactg catttgaccc aagtcctgaa gagccagctc ctctctcttc ctaattatta 96660
gaaggttttg tttggaccca gtgtttcacg tgtatacaat acaaacttct ctcttttcta 96720
cttggatcaa atttgttctc tcaaaataag attcccagca gtgagagaag acaagacaga 96780
gagatccaac atctctaaag ccatgaatca gataaccagc cacttgttct cttcagtgct 96840
gggaacagat acactgttaa ataaaatgat tttatagatt cttctcactg cctttccaag 96900
aaggggattt atcaacttca gggcacagca atcatttatt cccagactac tggcatgcat 96960
atatatatat atttacttct cttgacttag aaaaaagaga gaattggagt tgtgaatatt 97020
cctgtctccc tcaccccagc cccсttgaag tgagtcagga caaacttggg gcccaaatgg 97080
agctgtaagt aactgagtca catgcagaga tgaaaccttc acagacccac tgatatggag 97140
gttgaagatt aaattcccct ttgagaataa ctgggtaaca ctcatacaga gactactttc 97200
aagaaggcca gatcctccct ctaatgtata gtgcaacgtt cctaaccctc agcccactcc 97260
gtcatacccc cactcacatg aatacacaca taagcagtaa tataaagcac ttcccaccat 97320
agggcagcaa agaaggaggg aaatctttat tatggaagag tggaaggaag gaagggaagg 97380
gaagggaagg gaagggtaag aggaagaatt ctcagggtga gcagaggaat gacatgtttg 97440
gggcataatg aagataattg aagtgcagag tttgtatgga aaaatttgaa aatatcaggt 97500
ggcaggccag gcatggtagc tcatgcctgt aatcccagca ctttgggagg ccaaagcagg 97560
cggatcacct gaggtcacga gtttgagact agccgggcca acatggcaaa accccatctc 97620
gactaaaaat acaaaaatta gctgggttta gtggcgcatg cctgtaatcc cagctactcg 97680

ggaggctgag gcaggagaat catttgagcc tgggaggcaa aggttgcagt gagtcgagat
97740
catgctacta cacttcagcc tgggtgagag agctttcttt tttttctctc acaaaaaaag
97800
aaaagttcag gttgcagaga tggatggatg gatggatgga tggatggatg gacggataga
97860
tagacattac agagagtttc caattcttag gatgaattgg aatccttaag tctttattct
97920
gtaagaaagg aaggggagaa taaaattttg tgattttaaa atattttcta ccctgtagag
97980
ctaccctaca aggcatgaaa accttaaaaa aaaaggcatc tactttaaaa gaataatgtc
98040
taaaaaatta gaaattccct ctttttgccc tgacctttgg gaaacagagt gagtgatcct
98100
tttgaggttt ttggcactgc cttgcctgtg atcatatcct gaaccctagg tccataatca
98160
tgcagttacc tcagatgtcc ctttccctct agccacaggt aacacgctct ccaggcactg
98220
ggaaagtggg taattaggaa agcagaggag tacccatggg ctgtgatgcc cagttataaa
98280
cccagacatt tcagaattaa cagaatgagc atcaagtcct caaatgggtc tacatccata
98340
aacatgtcca gcagtcagct ctttactgtc agtagagaca aaatgttcct acactttccc
98400
taggggaagc cacatcctca gtaggttatc tctgatgagt ccagctagtc acaggtatgt
98460
agaagctgca tgcagcagag ggctcaaagg agggtccaga atagatacca aagcaaaagg
98520
ggagtctgtg cacgttctca cacgcacccc gaaacactct ttttgttcac aaaatagatg
98580
gtgtagggta gttccaagag atcatttagc tcaggttcct gcctccataa aataaataag
98640
ccttccatat tagttgtctg ttgctgtgta gcaaattgtc agaaacgtag aggcttaaag
98700
caatacccat ttattatctc gcaagttctg tatctcagaa gtccaggcag gcttgactgg
98760
gttctctgtc caagttctcg tgagactgaa atcaaggtgt tggccaggct gggatcttat
98820
ctggaggctc tgaggacata tacgcttcca accttattca ggccatcagc agaatcccgt
98880
ctcttgtggc ttgaggttgg aggtccccgt ttccttgctg gctgtcatcc agggaccact
98940
ctttgcacct acaggctgcc tatgttccta ttcacaagac accgttcatc ttcaaaccaa
99000
agcagcatgt agaatctttc ttgtggctcg tatctttctg gctttccctt cttctttagc
99060
cagagaaagt tctttgcttt taagcgttca tgcgattcaa tcaggcccac ctggataatg
99120
tccctatttt aaaggtaact gtgataccgt ataacatttc aggagtgata acagcacatt
99180
tacaggttcc aaggattggg gcagaacatc tttgggggaa cattttagaa actctgcctc
99240
cccactcacc cataatcctt ttaaaaacca aatcttgaag cctttttttc ccaaaggcct
99300
ttttgaataa gcacatttat acctaacttc atcagacacc cactttgagc aaacactagc
99360
atgtggcaaa ataggctgta aatcaatcag aactattctt tcccaccaca atctttctca
99420
aacacattgg gagaatctga cactgtcagt ggtataccag agcagactcc taccatctca
99480

caagagctga ctgttaaatg tttagtaatt gtggacattg gttgttaaac tattagtagc
99540
ctgaaattga ctatagtgag agtattttca ccatggaaag caaccgttcc aaatcagggt
99600
ttctctttat tcctgggaag ctggtttatt agctcaccac tggctgtagt cctttagggg
99660
tcattacttg acctcctgta gcatgcagga atcctctcca tggccttttt tatgcatgga
99720
catcatccta ttttttaata ccaggaatgg ggtgatcact ctcttataag ctagttcatc
99780
tccctgatgg aatggtatgt ggtagagttg aaacccacct ccctggaact tcccaccaac
99840
ttcctttgga agcagcactt gtgacagccc cagaaccatt tggagtaagt agcatttcct
99900
ccaggagaca tctctcctct ggatccacaa atcaatagtt agatgcaaaa tctttagagc
99960
cacactgttt gaattcaatt cccagctctg ccacttattt agttataacc ttaggcaagt
100020
ctcttaactt ttctggtcct ctggttcttc atgtgtggga atggggataa aaatagcacc
100080
tacctcatag gttattatga atattaaatg agataatgtg tgcagagaaa atagcacctg
100140
gtctggcctc tacctatcta acaggttagt tgtgaggatt aaattactta atataagcaa
100200
aatgcttaga gctctgccta gcacaaaata agcactatgt aactattggt aagttaattt
100260
gaaatgtggt ttctagatct ctcatcatcc tagtcaccct actctggatg tactccaaag
100320
tccctctcaa gatatagtgt cagaattgac ctaattagtc cagcatttga ctgaaacgct
100380
agactttgac tccagccccc catccttgac tggcactagc attcaagccg cttctcctct
100440
ttccctgggt ctttaataga gtcagagcga cttctccagg ggatcttttg gccatggacc
100500
agtagcatcc acacacgctg gggccttgtt aaaaaggcag gctctcaggc cccaccccag
100560
atctactgaa tcagaatcca cacattaaca agatgcttgg gtgattcatg tgcacattaa
100620
agtttgagaa gcaccgcttt cagggacgag atgacacact tattttaaag agaacgccaa
100680
ttagagaccc taagccttct catggaacag gggccttccc ctcagacctt gggagagggg
100740
tcagggaaat atcagtgttg ggttgttggt gacaggtggc ggtgggggt tcagtccacg
100800
ttcaaagagc cagaaacctg gcaggggaag agatggggca gtgacaccca accggaaaaa
100860
taaaggaaac tacaagaaga acccagctaa gagatgtgag gcttctgaaa gctcccatgg
100920
aaaggttcgc agctcctcca cctgctcggt ccagctgccc caggtcaagg aagctctgtg
100980
agtgttagct gacccggagc agcaaggata cattcagaag tgatgaaagg gaacgcttct
101040
tgacagggta aagagtcatt cagtaggaat gagacaggaa gaggtcacag agtcagaagc
101100
ccagcctgta ctcagagatt atttctggca tgggagggcc gaagggttag gaggccacct
101160
actcacaata caatacagag gcagatccac ttattacctg cctgtgctgc tgggatttca
101220
gtgtggaaat tctgtgcctc ctcactgtgg ctgcagcttg ggaatgacat ccagagctta
101280

cccacctgca taagaaataa gctataggtg taatagggg acataggcta aaatcctagc
101340
tcagctgctt aatagctgtg cgactgagca agttacttaa cctctttgag catctgtttt
101400
ctcatcttta aaatggaagt aatcataatt gaccaggccc agtggctcac acctataatc
101460
ccagcacctt ggaaggccga ggccagtgga ttgcttgagc ccaagagttt gagaccagca
101520
tggtgacacc tcgtctctag aaaaaataca aaaattagcc aggcatggtg gcaggtgcct
101580
gtagtcttag ctactcggta ggctgaggtg ggaagattat atgagcccgg gaggttgagg
101640
ctgtggtgag ccagattgtg ccactgcaat ctagcctgga gacagagtga gactgtgtct
101700
caaaaataaa taaataaaat aataatatct atgttaataa agcagaaata agaatgaaat
101760
aagaggcctg acatggtgac ttatgcctgt aatcccagca ctttgggagg tcaaggtgag
101820
aggatcactt gagcccagga gttcaagatc agcctgggca acttagtgag gtcccatctc
101880
taccaataat aattttttaa aaattagctg ggcatggtgg catgcacccg tggccccagc
101940
tactcaagag gctgaggcag gaggacggcc tgagcacagg agttgaggct gcagtgagtc
102000
atgatcacac cactgcactc cggcctgggt gacagagtga gaccctgtct caataaataa
102060
ataagaagaa tgaaacaaga aagttcttct tatggttctc atggtggtga gcacaatgta
102120
agcatatata ttatcttaga attcttcctt cctgtataaa gaaggcctcc tccaatgtat
102180
taatcatctg ttcaactaat aaatgctgct tactcccact ttcactctaa aggaactcaa
102240
tggctaaaga gaacccttcc cctttgcagc accctgagga tcagaggcct gatttgaatg
102300
tcctcgatgc aaaggactat ttcaaaaggc cagccaggca gcccagacat gtatttccta
102360
atcgtctcca ggttgtttga tagaagatct cctgggagca ggtttccgca gcagctcagc
102420
caggtctgtt ctgggaacgc tgtgtgcatt ggcacctccc ttggcagaaa gcttggagga
102480
aaggcaggtg caggtcctgg agcctctgac agcattactg gctctaggag tagctgctca
102540
ggataatctg tccccatgac cattaagtaa ctgccactgt gcgggaagaa gaactggaaa
102600
tggggggccc aaaaaaatct gaaaaccctc acttgaacca gtaagttata ccctgggttg
102660
ctgttggaga gagcttcctt ggagtagaca aatgtggtat gttaagtaaa ctggggatct
102720
aggtttgatg atactgggtc tgcagcttct ttgtcccact gaaaatcctc gggcattcca
102780
tgaaagtagc cttcaaaata tttttgtctc taatgacata tttttgctgc aaaaagatga
102840
gtggattcat tttacgaagt ctcaagtgtg ttagaaattc accatgagtc actcagcaag
102900
ttatgtttga gggcgttctg tatgccaggc actgtgctgg gcactgggac tactgtagca
102960
agtcagatag acaagaactt gcttgatctt ggaagtaagc agggtggggt ctggttagtc
103020
cttgaattgg agactgcctg gagatactgg atgctgcaag cttttgaaaa aagacaagtt
103080

ctctgtactt gcagagctta catccagtaa ctaactaact aacttcaggc tgtgttgagt
103140
gactgaaagt ggtggagcca ggagtcctct agataaggta gccatggaag gcctctccga
103200
agaggtgata agtttactca gagacgcaaa cgatcaggat aagcacagac cccggtgaag
103260
agcgtcccag gcagagggga tagcaagggg attgccctta ggtgggaaag ggcttgattt
103320
gaggactggg aagaccagtg tgtctaggac acataagcaa ggggaggacg ttatgaacga
103380
ggtctgaggg gtcagcagcg actggatcat gcaagctccc ataggccatg gtaagggctc
103440
tgtgtgtact acaattacag gatgcatgat aggacctggg ctgcattttt aatagttaac
103500
cctggctata atgtggggaa gggattgaag aaagagggca aaggcaggaa caggaaaatc
103560
tcttaggagg ctactgcaaa gcccaaggga gaggtgatgg tgttttgttg ttgttgttgt
103620
ttgttttgtt ttgctttgag aaggagtctc actctgtcgc ccaggctgga gtgcaatggc
103680
acaatctcgg ctcactgcaa cctccgcctc ttgggttcaa gcaattctcc tgcctcagtc
103740
tcccaagtag ctgggattac aggcatgcac caccatggct ggctaatttt tgtattttta
103800
gtagagacag agtttcccca tgttggtcag gctggtcttg agctcctgac ctcaagcgat
103860
ccacccgcct cggccttcca aagcactggg attacaggtg tgaggcaccg cgctggccaa
103920
atgatggtgt tttgatctgg gtcttaaagg cagaaggaag gggggtagta aattaactgt
103980
gctggggaag agagggaggc ctgagagtga ggaaagaatg aggggtgatt ccaggtttag
104040
gaaaactggg caatttgtta gatgatggtg ccattgacag aaatgggaaa gaacaagttt
104100
ggaaagaaaa ctcaagatct ggctggtgac ttgtattaaa cttaaagcct catttgtgac
104160
ttgagcagaa gtaaggactt tctccagtgt tcaagagctg gaagggattt ttctagcctc
104220
caggcaaggt aataccataa gtcccaacag tgatgccctc cctgggaatg atctcaatgg
104280
gagaatccta taccctgcct cctccattca ttccttgctc tgatggtggt tctggctggc
104340
taacctaagt tactcttgcc actagttaac gcctgtcctt atttctcttg tccccaccta
104400
agatgtcaat caaaacagca cgagccatgc tatgtcacat gacatgttgt ctgtccagcc
104460
cagagcttgt tgctgatggg ggcacagact agattttgag agaaatctct ctgttaccac
104520
ccttaacatt ccaacccccct ctaatagccc atttaggatt tatcatactg tttcatccaa
104580
acctttcatg acctgatttc tatttccagc ttcaaccacc ccttgggtca ccacctgtac
104640
ttattgagtt tccctagttt tctgaattaa tgactgaaga tgataagctt cccttacata
104700
tgactctcaa accaccaaac tgggattgtt gttactctta gtgataatgg ttgctattta
104760
tgaaactttt aatagggaac acaaaccctg cccagaaatt catataaatt atttcattta
104820
agaacatcac aaagtaggtg ctattatttg accttacacg tgagacttga agaactttag
104880

agcattgccc aaggtcaccc agctagtgag gggtggaggc gggatttgaa tccagctcat
104940
ctgtctccat tacctggaag aaggaaggcc agagcatcat ggcctttcac aagttgaaga
105000
gccacgggct ttctacggta gccagccacg cttttccatg actggggtgg gtgtggcaag
105060
tgatgagggt ttggagttca tgtggtgggg tggcagggac caggtgtctt ggtaactgct
105120
gttgcattca cttcaggagc aaaggaccag atctgattct gcaggatcaa caatatggac
105180
actgcaggct ctgtagacat ccaaagctct aatggtgact tggggaagct caggagggca
105240
gggaggttgt acccatttag aatgtaaaga ttcctatttt ataaaaaaga aaaaaaggag
105300
actgaaggcc tcagtctcct ccaacaaagc caggctgtgg ggtagcagag tctcaaaggg
105360
tgcaggccca tggccactgc ccagggctcc tgctcaggcc tcctcactcc cacaactgag
105420
gggagaccca gttccacacc cacccaccta gcagtgtctc acacccaccg ggagaggtct
105480
aaacatcttc cctgggaaat ggtcccaaaa tgtccctgca gtaagcaacc atctggagag
105540
gcccaggtct acatctgttt ttaaagctcc aataaataaa taaatgaagg aagaaaaaaa
105600
gaagaagaaa tgcagaacag ggtgactaaa attggcatgt atttttaaat gtttatatta
105660
acaaactaac acctttttaac atgaaaagca atataattgt gctagccaca aaatcatcgt
105720
aggactgaga aaggaatcgt gattctgaga gccctagagt taatgtgatc cagctggctc
105780
atccctgtga ctgcagaagc ctgtttggag atagtgtcag tagcttttca ggccctctgt
105840
gaattgccag aatgtgtgac atgagccaaa tttcccccca gcatccccgc cgccgccacc
105900
accacccccg acccaaccct cccgccggct cccatagaat agtcactgcc atacagaaaa
105960
agagaagttc tactatttct gggcaagatt tccacaaacc agtttgtccc tttctgcttt
106020
catgaaataa accatttgga tcaacgtcag ctgattgcaa aaattttccc ttgtctcaaa
106080
agcaagactg ataaggaagc aaacatggga ggaccttagt ggccgagcct ttatgtgtat
106140
gttatttcat tgctctcata actgccctgg gatgctgtaa gcatgattca tcctgtttgt
106200
ttatcagtta aattatgtat ccaagattac acagcctatc caggattaga actcagagcc
106260
ctcggctgtg aagcttgagc tctttctttt cagtcttcaa atatgatcat gccatgaagc
106320
agcacaaagc ccaggaggag cccagtgagg ctggaggggt ccactggcag ccactctcct
106380
ccgtgcccct gtggtgttgg ggcaaacttg gatctttctg aatcttttaa ctgtttcctt
106440
ctcttcccgt ttttgtctgc tggctgactt gtcctacact ctactccttg cttatgatac
106500
ttatttttcc atccacagca aaacaattca catcaaggta attgatgatg aggcatatga
106560
gaaaaacaag aattacttca ttgagatgat gggcccccgc atggtggata tgagttttca
106620
gaaaggtgta gtaccctgtc ctccacacta acactaacat tcttctctcc tcttctgttt
106680

cttcctctcc aacccatttg tctcctcctc ctcttgtctt ccacctctct ggttcccttt
106740
cccttgtctc ctctcttgct ctctctcctg ctctcttttc actcctccct ctcctctgtc
106800
ctctctctgc ccccagctct gtcctaacac ctgccagcct gacacatggc atccatacga
106860
gggatgctca agaccgatgg taattgttct gggataagga aatgagtatg gggaaagaaa
106920
gagccaaaat gctggagtat catgtgcggc tcttggcttc tccagaatgg ctgggcataa
106980
aggggggaaa agggaccaca tagcccagca ccagacagaa gagcagcact gagaaacagg
107040
ctttcagcac aaatttccat ggggcagtta ttctcagggc taaacttaga gtcccaggaa
107100
gttgagaatc aatgtatttg gattacagtt cattcccctc ccaaaagcag gctttaggag
107160
ccaccttatc tgccatgttg ctactatcaa gacttgtttc tcctcctgac cttgaggaag
107220
ctgaaagtac aggtttgagt tccagatcta ggtcaaatat ccatttgtct tcctatgttt
107280
ttcctattaa gaacacccag gtgtggaggc agagagttag aatagtggtg gagatcatcc
107340
tgacccaaat ggaagcttcc ccaagaggtc catggggctt ctcagagtgg atggaatctt
107400
tgccttcaac ttcaatgacc ccatacatcc catggcctcc aatagacaag tcaagaagtc
107460
ctttcctgaa tagatcatac tgtggagcag ggagctgcca gtactgaggg caatgttcct
107520
tccccttcca agctgtccct catgccctcc agtacatgcc tgttgtcaca gagcacccca
107580
atcccatccc acagcagagt tcctgcagca gagaaacagg ctcacacctt gtagacagcc
107640
ctggggtccc atatctaggg ccaacagaaa tattcccaaa aaaatgcctc ttgacaatca
107700
atgagctttc tcttttgtcc gctgagcaag gtataaaaag atgtcaaaag aagtacccaa
107760
aaaggtaata aaaatgtaca gtcgtgcatc acttagcaat aaggatacat tctgaggaag
107820
gtgtccttaa gcaattttgt catcgtggga aaattataga gtgtactttc acaaacctag
107880
atggtgtagc ctacaacaca cctggactat gtgggcctat tgctcctagg ctacaaacct
107940
gtacagcatg tgcttgtact gaatattgca ggcaactgta gcacaatggt atttgtgtat
108000
ctaaacacat ctagacatag aaaaggcaca gtaaaaatat cgtagtatat agccttatgg
108060
gaccactatt gtagatgtgg tctgtcattg agcaaaacgt ttttatgtag catgtgactg
108120
tacttgtaaa gtacacacac cacaaatgca cagcaagtcc tgtgccctac aagccccttt
108180
gggtcagtct actacattat aaatggcaaa gccgagcacg cccacagaag gtagcaggaa
108240
catcagagga tctgaagaga catttaggta aatgctcttt accctttaga gcatttagtt
108300
cttaggcctc ccctccccca atctcccccc cgcccccccgc caaaaagaaa aagaaaaaga
108360
aagcagaaaa ttacaattct ggctcactag taggacctgc tagccaccat tgtgattcca
108420
tgaaggacca gaagaaacca tataggaaga atcaggccca cacggcaacc tctccacatg
108480

acaaagagcc agtctttgga gggcagtgaa tttcaaggaa agttttcttc cctgggtgac
108540
ttgtttttaa aagatgttat gttttgttga gatacccaga gatgaacaga aacttccatc
108600
accttgtgcc ccagacccat gataattcac attgaggaaa ccagttttgg aacacatcac
108660
ccctaagtga tagaagccca aaggtgattt agaatttgat gatttacatc attttcttca
108720
cattttccca gaaatgcatc agctgtaaat agtaaaggat tcctatgtaa tattgtggtt
108780
aatacatatt tattttagtt cccaccactg aagccctatg agataaagaa tgagaaagat
108840
cacacaattc tacctccctt tcttctctct ctctctctct ttctctttct ctctcactct
108900
ctctctctct ctctcttctc ttcctctgtc tggttttcct tcctcataaa tacttttctt
108960
ttaaaatttt ctttctgaaa ctcacaatgg aagtgagtat agacataaag aagggacaca
109020
agccctgggt tctgttgaca tattccctgt tgtgggaaga ccctgggtta ttcccagtgg
109080
gttagtagtt tacctgttgc ccagagaaat gccactgtta tcatgtgaca cccagtggaa
109140
tgtgctgcct gactcacttc ctactaactg ttggcaaggt ctaaaatgac tcctcctcac
109200
cattacccgc cttctgcctt ctcctcccct tctgtccttc tggctccctt cctttgccca
109260
cctttccttg cctcctggct ccctgccccc tcacccgtaa gaacaactat gaccaagaag
109320
acaagaaaaa ctaagaccat ttattacctg agaacaacac aatccaccat ggtcctgttg
109380
aaagccacca tggtgggact ggactgcatg tgccaggaat gacggggaat gattttaaag
109440
gctgtgctcc aggtgaccaa ccaatctacc gacccagtcg acacactctc tctcttgttg
109500
tccctacagg aaaaccataa gggttaaaat agtagatgag gaggaatacg aaaggcaaga
109560
gaatttcttc attgcccttg gtgaaccgaa atggatggaa cgtggaatat caggtgtgag
109620
attctttaaa aacaaaacaa caaaaaaaaa gaaagaaaaa ttaaaacaaa ctgaaaaaca
109680
acaacaaaaa agaaaaagca gctatatttt tgtctccctc cttttcttcc cttctcctcc
109740
tttctctttt tgaccaatgg attttttat tcttttccct cctgtattct cgctctcacc
109800
ctgtttcggt atcatctctg ccttcttagc cttagcttat tccaaattcc tcctttaccg
109860
ccttctgggc agcactgcag cctcaactcc tcattaccct aatgagttat ttccctgttt
109920
tgctacaatt ttcaattatt caattgccat gggcccctgc actctccccc accccacccc
109980
tacactgtaa cctgtaaatg tgaaaattcc ttggtgggtg gggaggagaa gaaaaaaaag
110040
gaatgtgatg cgatgcatgc ctgtgcccct tcctgccttc ctcccctgcc accccctcact
110100
ctttagcctg gattgaatgt ggggggtct gggatggggg ttggggcctg ggttgcaatg
110160
atgctttgac agttttctgc tgcattcccc aacttccttt gaacgcttgg caggttattc
110220
acttgtggag tggcccatag gcccctctgc ccttcgagga ggtaagtgta ttttctggct
110280

gtttcacagt tgggcagacc gtggcatggg aaagtgtacc aattgtcaga agccacggct
110340
tctgagagct ctgagagaga gagttgactt ctggggtaat catgcaatct ggaattctga
110400
gctattcttc ctcctctggg catcccaccc catgccattc tatgttccta gcccaaggtt
110460
gggtgcctca ttcaggctac tttgggacaa tgcaacctct aaagcagaaa attgagagtt
110520
cctgaaggga aggaaatagt tccaggtatg aaaattcccg tagccagggg ccccagaaaa
110580
ggactgacat tgggcaggcc tggagtgttg acttgtggat tttccaacag aagagactct
110640
aaatgatgca gttggtgctg atccctgaca gacaggtgtt ggaaaggtca cagatgtctg
110700
cctttgcttg gcatctgcaa gagaaagtac cgcccagatc ccaagatagc cctcatccca
110760
cactagagaa gtggcctcat ctcctgcttt cctcaggacc tgcatctgag aatacctgcc
110820
aggggctcat ccctaaagga ctgattatgt tgcaaccagg gtagaagtaa ggaaggattt
110880
cttcccttga agaaaatgat tggaagccac tactttgaat ggcttccaat catttggagg
110940
catagatgtg ggaatgggtt agggtgctcc tgggaaataa caagaggacg ttcacactcc
111000
cattcaggag agatatgctg ctgggagcct cctagcaaat gaagcagtga aatccacctg
111060
tttgtcaaaa aggggtgatc atactgcaat tagttcatat tcatgtgaca aagagcagca
111120
taaaactttc cacacgagga cagagctaag agattcagca acaacattcc caaaggattc
111180
tctacaggcc ttctcagtgt gattggtcat ttctcattgt ctgctgggga ctctcctgca
111240
gagctgacca cttctgtgcc tgcgctggtt tggacacacc tgatgctcta ggggcagaac
111300
tcctctcctt cttcactgct ggttctcttc gtcaccactc aataaaacgt tgccctcagc
111360
ctgactgcca aaaagtgctg gaagaaagaa attatctctg gttctattgt ttcccacatt
111420
gtattcttgc ccaacttcca gttcttgcca ccaacaatat tctcagaggt tgcctcagca
111480
cctgccctac ctcattccca cctcccttga gcatttattc catgtattca taattggttg
111540
gaagcagcag atacccaagg ccaattgtaa gtcaccttca tcagtttcca cagtccaagc
111600
tacttagatg caaacgaaag cagcacatgt acagcgtaca ggaaggaagg cagtggttcc
111660
agacaagagg aagagattgg aagtccatac atgcctttat tccaccagta aaaaggctct
111720
tctcttatgc ctcccttaaa acctctacca acagcaggac agagagtgac ccaagataag
111780
tcttcaagag acctaaccaa atgcaaatgt ctttggctaa tccccattta aggacatctt
111840
cctgttttgc acagattctt tgcccaagga aatgtcagca atgccctcgt ggagggagta
111900
ggtgagaaga caaggatttc agcaagctat ctgtgtggtg tgccccccaga tctccccagt
111960
gaccgagatg ccaagatgaa gagtgccaag aagaaattgg tcaattttcc agctgcctat
112020
tttattgtct atgttttcta ggcggttaat ttccagtttc ttcagtactt cccgtatttt
112080

gacattagac cataaggtga aaggtcataa aacctgattg tctagactca gaagcaaatg
112140
gaaacccatc caaatttcca gaattccctg ctgttctcag agtgagaaac agaacagtgg
112200
aaattgcttt tcattatcac tactgcatgg gagagtctga aacattcaga atggcatagt
112260
ctttgcatgg tcaaaatgac aattgcatta aaaaaatgag agactggatt tgaaatagga
112320
gactctattt ttggcaaaca aaacagactt cagagttgag attaaaagct ctggatgagc
112380
tgggggatgg aaaaaaggga aggaaaaaag ggagactgaa taggaaacac agttgctctg
112440
gagtctagaa gtggacttcc gagagcaaca ctgagcaaca taatcaagac tgttgggcct
112500
gggcctggac attggaagcc ttcggataga aaggaaagct ctctgtctct ctctctctct
112560
ctgaagaatg gggcctgttt ggtcctcctt tttcgacaac cgtgggctca tcttgacaag
112620
ctgcccagat gcttcctaat tactcacagt cctatgctct ttccagcttg tccctggggt
112680
gtctgagcag gaataaatga ctctcacctg acccagggga tcaatacagg ggaaagttca
112740
gctccagctt ctctcatgag cagcagcagg aaaaacaccc tcgaggtatt gtgtcagtca
112800
aagctggcct acccaggtct tgctgaccca tctataactg ctgagcagaa agtcttggat
112860
tcatggagac aatgaccaga gaatgatgga attccagcca actgcaggcc ttctcactac
112920
tctagggatg ggccagatgt tcggtggcat gtatgagtga aaaccagggc atcagggacc
112980
tttctggaag agctgccttt gtctgaccca cctgtgttca tttatgtgct gggatctctg
113040
atctcccctg gaacttgggg gaagctcttc cacgcaaact cccggaagga gcagaataaa
113100
caagctcttg cctatctatc tatctatcta tctatctatc tatctatcta tctatctacc
113160
tatctgccta tctatatcta tctatctcaa tgtagtgagg aaagccattg atccattaac
113220
ctttggaatt ctacatggga gatacctaaa aaagtgaact gccttgttta tgtatcatgc
113280
agactctgga tccacatata tctcagtggc tgtgaatata ggatgattga tcacaggcct
113340
gagttgcatt cctacagatt cttaggaaaa aaattgattc acagacatgt cccccctggt
113400
tcccccacaa cacacactcc ttcctcagca atctctatca gtcaccaact acacgttgaa
113460
tatgtggcaa gctcttccca gacctttatc tgagagccaa ggagtgaggg gctgtactaa
113520
gatatcatag aaatgaaaat gtggtgtgtc acaagtttcc ttaattctta gatcttaaac
113580
tctaagaggg ttcagcataa gtacaaattc aagggctaga gacaacctgt attgggtgtg
113640
tctttaactc agtttcccaa tccacatagg gaccttgcat ttgtcatctc tcatctatgt
113700
atagctgttg gtatgacagt ttctctgttc cagaatacct gaactctgac ttagcctgtc
113760
ctttctgaaa cagaaaaatc acccaaccag agatctatga gatctatgga aaagacagtt
113820
gccaaaatag acagcaaaca gccaaactta attgaacact accacatgca gggactttgc
113880

taagcagagg tgatacaaaa tgggaggagc ccatagccct aacttccagg atatatctac
113940
ggtaaagaca aaccattcaa ggaaaacatt ctgcaggact taccttttg ctaagtcatt
114000
cttttagggg aaatcaaagt tctagtcaac gtggcagcta ggaaggcatt tgtggtgatg
114060
gaaaccttat gagcactgag aagctgagca tgagttcagc taagtcgtta gggatggaag
114120
acatagacct gggcactgtt ccactcttgc acaatgctac ccatttcctt gagctcccat
114180
tcaagcccca tggtcatttt tgccactcat aagttagcta ctctggcagg gttgcaactt
114240
acacagtttt catgataact ggattctcac tcctttttt acagaatgga tgtgataacc
114300
tggtatccta cacagtcatg agtgaccaac ctacccattt ggttccccat cctcattcct
114360
ccattcctag ccctagggta gccgggaaag cataggagca aatgccctta ccagggccct
114420
ggtgctcagc agcctctccg gctgctcaca cctcttgctg ctgctctgtg catgctccaa
114480
aggctgcttt ttgcgtatgg ctgctgagct ctcacctact aagctctctg ctttccttat
114540
gctgccagca accacaaaac ctggtgatac tttcaagatg ggacattaat gctctttcct
114600
tttctttctt ccatttttct ggtatccatt tgcaaacagc gctcctgtta tctccaggta
114660
agaggtgtct tgtccccctc ttttctttcc acttcttgcc agtgccatta tttggtttaa
114720
gaccaatgtc ctttgattta ttgaataaga actgcaggct caagttaacc tgacaatttc
114780
tcccaaggac tgggagattt attttcccac atgaagcaat tatgagaaag caattgtgag
114840
gaaggcaatt ccttgagcat cacttctgtc tggggacgtg ggttaaggca tagctgatcc
114900
tctctgggac caggaagaga aattaagctt aacaaggaga tggtgggtca tagacttctc
114960
ctgagtctta attcatctgc catctcatgt tgtgggggaa gagacagtga gattcagagc
115020
tggaatctcc taatataatt gtgacaggat ttgaaaaaaa aatactttaa tcccaaggga
115080
tccaggaaat aaccaaacct gttgtgagaa taggaaatgc aatttttaaa gaatctggaa
115140
ttttaccagt cctggagatc ttccatctca tcacagctga gacttaaatt gctagaattt
115200
tggttcattt gtcattgacc cttaaagtcc tatgtgccgt gaacaagatg aattaggatg
115260
ggggattggg gcagtgttct ggctggaaat ataaatttta gagaatttat tttgaagaga
115320
ttctcatgca gaatctaggt gctatagagg acgtacacct actttgagag tatgcttgca
115380
tgagtggaaa ccaatcataa acaacattca acttcatgag cagatatgaa agcattttca
115440
gcatatctag caatactata actctttgtg caagcagagt ggcctacaca agacagtttc
115500
aatatatttt aaaagaacgt cttacatttc atcagtcctt tgaacacaga aaaaaatgtt
115560
aaggccactt aagaggcaaa acatcttaca gagttcattg atattcaaag tcacctacag
115620
gctacatctt gggttcagga aggggcggtg tacatagtaa ggacatacgc cttctgggag
115680

ccttaaacaa acaaaaaaaa tgtaggtaac tcctacattt ttcttttgtg gaaaaaacac
115740
agttactcca gcttccttgg ctttttgctt cttttttata ccaacaaaat aagggctatc
115800
ctcaaccctc tgttcttcat tcttctccca gggtattgat ttcataacat tgggtttttc
115860
ttctctactt cactcatcct cttgcctgtg aaggtatgta aggcttcttt gttccaactc
115920
tttcctccac ccgccccccc tcacataaat gcataacaaa gattgtgatt taatttaagt
115980
ttctttctac ttttaacata tttgcaaaca tcaatagaag ctaaaatggg aaaaaggaaa
116040
tgtttctttt cctagctctt tcaatctgta agcctttaat ttaggagcgc tgattagcct
116100
ttcaattcgt tggaaatctc aaatactggt tttaattttc ctaggtggac agagacagag
116160
ggaatatgtt cattctgagc taaccacccc cccacccccа agctcaggcg ccttgcagga
116220
agagcactag ctacatcact ctgcagagtg ttcacaacat cctattcttg tctggcctgg
116280
caagctcttt gtccttccaa tatttgttca atcttccatc ctattcatat tctatctttc
116340
tctcccctcc cagcctctct tcctgttcct agaactgaga gtttatttag tcagtctgaa
116400
tatctagatc acctgccatt tattctcttt acttgaaatt ctgaggagtc acataaacaa
116460
gatatcagaa tcactatggt cctctaaatt gaagacttat aattctctca agaaattaac
116520
aacatttgaa tttaaaggaa agatcatgac aaaaatagaa aaaggcagga attattgcca
116580
aaccgagaaa ctagaaacta gaattaactt aaaggcatgt gactcaatca attaacaaat
116640
atatacagag agcctctgtg ggactgtggg agatccaaag atagaggatt ggttatttgt
116700
caaagggatt tttgcagaaa gctagatgga aaaactgact gtcaccacag aggtggacag
116760
gtcagtaagt agatcaatat cctgccagat ggatatagtg ctagattgat aggtagacaa
116820
ggggttagac aggtacattt atatgtcact ggagagctca ttatattggt ataaagttat
116880
tgtgtcacat gtaaagtatg acatggggga attggggagg aaggagtgga ataatactgt
116940
cgctgctaag ataggcattg tgatatggtg cttaaacctg caagtaaagg aaaagagtat
117000
ggaatctgtg tgtctttttc taagggcttt ttcccagagt agcttgcagt ctggcttcta
117060
gggttgctgg cctatagcca gaaccctaga ttcacccaga tttaccttca gaattaacta
117120
atcagagact caaattcaat agactaaatg aagtcaggct gctagaggat gtctgctgac
117180
ttggacatat gcagaaagac atggatcctt gagaaaacat tgtttccaaa agtggccacc
117240
agcactagag gaaggacagc accacggaca gctcccagac attttaggat tgccttctgt
117300
gtttggtgcc cgaacactga gcaaaacagc gaactcagga agtctccaca cactctcata
117360
ccatcttcat gcagtccaac taagaaaatt cttacataaa atataaggct gtctgcttgg
117420
117420
taatttaaac ccttggctta tagtcttttc agtgaatttc tttccttgca aactcgagag
117480

ttggagtctc acgactgccc ttgcttcacc aattccccag ctagagacaa aagaccttct
117540
tggcctctga cccattttgt ccttgagatt atccaaggac tacaggattc ccctaggagg
117600
tttactgtgt ggaatgaaag caattaagga gctgaataaa agaaataatt gcatgtgaga
117660
atgtggactt ggatgggaag atgtttaaat gagctctgaa agaaacaagc tgccaagagc
117720
aatttctaa ttaaagggga ataaaaagat tcaatctcta tttcactcta atccagaaaa
117780
catgtcttca tggagaagtg ctcttaaaat ggactcatca gccaaagtgg aaaaacaaaa
117840
aacaaaaaaa ctgttcaaca tgagaaggga ccattggtaa atgagtcaag atgctgtgaa
117900
accagtagac atttcctttg aataaatgta cttctgcacc ttcaagaact cttacaggaa
117960
gtggttgaac aaacaggccc aaaagttcaa aatagttcaa ggtcaaaaca cttgcccttt
118020
cttcccagtt ccccaacatc tcactgagtg tcttgagaac ttcacttgat gctatttctc
118080
aggagatgtt taggtcaggt tgtccaccca ggtataaaag agaaagagga acgcttatcc
118140
cagtctgcaa ggcacattct catggtctgg ttataaagtg tttagtactt cataaaaaag
118200
gcactaaaaa tatatataaa ctccccattc ccaagagtta tttgctttgt acccactgcc
118260
catgcctaat actctgagct gtatccttcc agggaatgga aaaggtgtta aagcgagtct
118320
gattttgttt tgttgcagat gtgacagaca ggaagctgac tatggaagaa gaggaggcca
118380
agaggatagc agagatggga aagccagtat tgggtgaaca ccccaaacta gaagtcatca
118440
ttgaagagtc ctatgagttc aaggtcaggc aaacagtgag gtctaattga ataataaata
118500
aattaaagtg ggaggcagaa gacctggggt gtttttttcc actttcacta gtgaatatgt
118560
gaagttgaaa ctgaacaaat cacttaccca ccccaggtct cagtttcccc atttgtaaca
118620
tgaaacaaat agtgctgacc atttgtatgc taggaatatt gttaggaaac ataatataga
118680
atgtgaaata agtggactag aaagtcctga gatgtattat cattattgtt taactgtgtt
118740
tttaaagcaa aaatattaaa actcactact acagggcaag atatattaac atcattatta
118800
ttattcatta ttgtattatt ctaaatagcc aatttcaaaa gtcacaacca ggccaggcag
118860
tgagggactc acgcctgtaa tctcagcact ttgagaggcc gagatggaag ggtcacttat
118920
acctaggaat ttgagaccag cctgggcaac atagggagac tccatctcta taaaaaataa
118980
aacaaaataa aaatcagctc agtgtggttg tacatgcctg tggtcccagc tactcaggag
119040
gctgaggtgg gaggatggct tgagcccagg aggttgaggt tgcaatgagc catgattgca
119100
ccactgcact ccagcctggg tgacaaagtg agaccctgtc tcaaacaaaa caaaacaaaa
119160
agattacaac caaaaacaaa gggaaataga aggattgcct caaaagagat cgcccaaggc
119220
cattccatgc gtaactgtca gaacaccttg gagacagggc atctttcatt cctttgaaga
119280

accagactcc tcattggttc tgagcattct aacctcatgg ttccaagttt ttctcttctt
119340
aacagactac ggtggacaaa ctgatcaaga agacaaacct ggccttggtt gtggggaccc
119400
attcctggag ggaccagttc atggaggcca tcaccgtcag tgcaggtgag aagtgtctca
119460
ggctggcctt gctgggagaa gcaggcaacc tctgagaagg aagcgtaaag ccacgttaac
119520
agcctgccag tccctaggaa ggcttgtgtg ttcagtcttc ccagctctgg tcctaggtgc
119580
ctgcttggaa aagaatcatg gcgtatctga aaaacatggt tatctctggt ttcaaatcgt
119640
tgttctgctg tgtgaactgg aacaatgtac cctctctgac ctcaatgtcc tctttccaaa
119700
ggggaactat tgctaccttt ctcagaaaag tagaaaggta cagagtcttg tataaaatcc
119760
aaactcaata aattctgatt tctgtcattc tttcttttca tgggtttggt cccgctcttc
119820
tgtaaaatgt gggacaattc tgatttagag atgtgggagt taggagttta taaaatgtgt
119880
tgcattgact ctccaacaaa acactctgga tgattccata cccctccctc ggcatttact
119940
gacaggctcc ctcagtagtg acccacagca cagccgggag tcctagcagc ctgaggggac
120000
tgctggttgg aacagggacg gaaaaggtct cccaaccacc atcactatca cctctcagca
120060
ccactgaggc ctcctggcct tgtcttttat tgagagactt tgttgtcata gcaacccaca
120120
gggtcatatc cccaaggccc cagagccaga gcaaaaagac agccaggaag agaggtttgc
120180
tgctgctgct gctgctgcta ccccactttt ctcatcacct gctttagatc tttctagctc
120240
cccctctgat gacctgactg tgcccctcaa gacaataaac ggaatgtagg ccacatcatc
120300
taccctgctc cttttacaaa ggaggggact gaggttcaga aataagagat gatttacccc
120360
agcttacaga ttttcttcat ggcaaagctg gaatgagaac ccaagtgttc tgactcctgt
120420
tctttcaaaa cccagcttct accggttatg ccaaaacatg acagaagttg ccgttggcaa
120480
ggcacaggca tgcctcagca taccctcccc tccagggctg ctgagtgggc aactctgccc
120540
acatttcctg gcaaggacaa tcaaggccca tcctgctttt tcccatgaga tgtttggagg
120600
agggcactgg ctctgcagta tattctcgtg atctggaatg acagccatcc ctcaggggac
120660
agataatgac cagaaccaca atggttattg cagcagtcag gtcagaaaat ttgagaggag
120720
ccctgctggc atccagtgaa gagtggccac accgaactga tttcacttct ctccttagac
120780
aacaaaatgc agcctgtgca ttctcctttc tttttttttt taattatact ttaagttctg
120840
gggtacatgt gcagaacata gagttttgtt acataggtat acacgtgcca tggcggtttg
120900
ctgcacccat caacccgtca tctacattag gtatttctcc taatgctatc cctcccctat
120960
ccctcacccc tgacaggctc cagtgtgtga tgttcctctc cctgtgtcca tgtgttctca
121020
ttgttcaact cccacttatg agtgagaaca tgcagtgttt ggttttctgt tcttgtgtta
121080

gtttgctgag aatgatggtt tgcatcctcc tttctttctg ctccactgtc ttgtccctct
121140
taatctcctt ctttcttctc ttccttattc cctggccctc tctctcccac tctaccttgg
121200
tgccctgcat tcaaattgac ctatgaggca gcccaaattg tttccccact attttctggc
121260
acgctggccc tggcccccac cagctgccca gaagacagct ggagtcccct tctagcggat
121320
gatgcctgtg gtgcgggttg ggcttgactt tctcatgaat gattatctga cttcttaccc
121380
gttctcttgc ctgtttatct tgccttcagc aggggatgag gatgaggatg aatccgggga
121440
ggagaggctg ccctcctgct ttgactacgt catgcacttc ctgactgtct tctggaaggt
121500
gctgtttgcc tgtgtgcccc ccacagagta ctgccacggc tgggcctgct tcgccgtctc
121560
catcctcatc attggcatgc tcaccgccat cattggggac ctggcctcgc acttcggctg
121620
caccattggt ctcaaagatt cagtcacagc tgttgttttc gtggcatttg gcacctctgt
121680
cccaggtgag agtgagaggt gcttgaattt gcaaagagga ttttacctgg ttcaaatgac
121740
ccctggactc catctcatta tcttccacac catctcagat ctgaacttaa cagagcctct
121800
gcccttaaag tgcacaaaag tcaatcaaag agatgaataa tgacattagt aatgacagct
121860
aatatttctt gagcactttc aatgtgacag acaccatgtg tgttcagcaa tttacacatt
121920
tacattttcc ccctgtaatg tttcccaaag ccctattaaa tagggtaagt tattatcccc
121980
acttcacaga caaagaaact gaggcccaca gaggttaagc tacatgccca agtaagtggt
122040
ccaatttctt aacctccaca ttatgtgagt agaccacaaa cagtgaaatt aaaagaatgt
122100
agatattgtt ctccttctat ttacctctgg cgatctctga gaggttaaag attagccagc
122160
tcaaagatat caaaggagaa atgcccacat acattcttgg cctcctctac ttggaaggac
122220
actgtgagta caaagtatct cctagcagga cagccaaagg aagttccaca gcttttatct
122280
ttttatagga tgaattacat actctttctt tttcttagga acactcagag acaaacagaa
122340
aggagcggac attcctttac tcattgaaca aatatttact gagcacctat tatgcctgtt
122400
acagtattgt gctagttttt gggactatag tgaaaggcaa gatacacatg cttccttctc
122460
cacgtggagt ttataatcta ctgaaggagg caactctcaa ctactgtaat taaagttatc
122520
ttgttaaatc ctaggaagaa aaagaaaagg tactgcatac ggaaggaagt tgggcctgaa
122580
tgtaggagtt agcaggtaga caggggctgc actagcccag gttctttact taattcagtt
122640
aggggctttg gggcctctga actctgaact tctgccaggg agctggcatc ccagttgccc
122700
cagaaagaaa cagagcacat cctcctgcag ggaagttagg ctgaatctca tcagacagga
122760
cttttctggc tgggccaagg gaaatctttc ctgtaccaag caaacatatc cttcaagaga
122820
gtagctgaat tcacatcaaa ttctaggaaa acctctttcc aaaaccccag cgcaggccag
122880

cggtattatt tgtccattag tgatgcaaga gatttagcta tcgtggaaat gcatcagaag
122940
gttggaaatt agatggatga tcccaggaag gcctgtggat gagatgccct gtgatctctg
123000
ttctccaagc cttgggggac ctgaactatc agaggggagg gaggaaatat gggggaaagc
123060
atagaggtgg gaagaaatat cagaggatca gaagcaaaaa acaacaataa caacagaaac
123120
aaaaacaaac aaacaaacaa aaaaacaagg ccataggcaa gaaagggtaa gaggttttct
123180
ctgggagatc taaaaaaaat ggcaataatg aggtaagcca ggcagatacc tttgggcatc
123240
tccaagtcct tgcaattggc caagacaaca gctaacaaca tttgaggctt taagaaggtt
123300
accctgtgat ccactcatct gatttagtgg ctttggctga agctctttgg atatagttga
123360
aggtacggaa agggtcctta catgaggact ttagggtcaa gtctcttgct aacatcctat
123420
gtgaccttgg gtaaattctt tgacccttat ttttcttacc tgtaaaataa aagaattggg
123480
ctagatgtct ctgacagtcc tccctgtatc tacaatctgt gccaagatct aaagtcaaac
123540
accctgcaag gccctgtgat acatatataa accacaaaga cagagccccg tcttccttga
123600
gtccacagtt caccctgcat gtccccatca tggttcccca acatgtcctc tgtccccaaa
123660
atccagcacc tcacccagtg ctcaatcagt aggcattgct caataactgt tggtggttcg
123720
tgaataaatg ccccatatga cagttaaaat caggcatcta ctccaagcag cttcccaggg
123780
tgtcaaggtt ccctggggag atattatggg atggcaaact tcccttactg aaaaagtagt
123840
caaaggagaa caataagccc actcagtaaa tatcagaact ggaaagccct tcagaatctt
123900
tcagatcact gcagatgagg aatgggaagc ccagactagg gatgtgacct acccagggcc
123960
acacggcttg cttgcggcag aactaggagt taggagtggc cccctagccc ttgtctctca
124020
ttcctgggtt cagcccacca gctcaagctg ctttttgggc atactggaag acaagccctg
124080
cacaccttag cctcctacca gttcccatgt gtctttgtcc ttttccagat acgtttgcca
124140
gcaaagctgc tgccctccag gatgtatatg cagacgcctc cattggcaac gtgacgggca
124200
gcaacgccgt caatgtcttc ctgggcatcg gcctggcctg gtccgtggcc gccatctact
124260
gggctctgca gggacaggag ttccacgtgt cggccggcac actggccttc tccgtcaccc
124320
tcttcaccat ctttgcattt gtctgcatca gcgtgctctt gtaccgaagg cggccgcacc
124380
tgggaggggа gcttggtggc ccccgtggct gcaagctcgc cacaacatgg ctctttgtga
124440
gcctgtggct cctctacata ctctttgcca cactagaggc ctattgctac atcaaggggt
124500
tctaagccac acaacagagc ctccagcagg gcaggcctag gacttctcct aagagaaggg
124560
cacttcccca ccagtgatct ctcccgactg cactgccctg gagaggcagc atcaggacct
124620
aagccccagg aacttcaccc aacttaggcc ctggcaatta actgaaaggg caaagtctta
124680

atcaatcaaa caatggagga atcaccgact ttacacagta tttaattgaa tacaaacaag
124740
caacagcaac aaatccacct ccaccccatc tccccctcat atccctgacc caaagcaaag
124800
gtcagagcct ttcgcctcct tctattccat cttttgatta ttcctttgcc tctcatttct
124860
ttggaagcag ggtttctcct ctctgcccaa ttccatatgt ccctattatc tcactcagct
124920
gacaagacgt gaaaatgagt cacattcatg tggctggggt ggggttcttt tttcattgta
124980
atcattattg tggttgcttt cgttttgccg ttaggttttg cttattattt tgttttgtct
125040
tttttttctg aagtgagtga aaaaggtgcc acaaaggaat tccaggtccg agccaacaga
125100
gagaaacatg aatttttaga cacatgctct cctgccacct cttggctcca tcaagatcca
125160
gttccccatc tcactgtttt ctctgagttc ttgggaggag tgatggtgtt ggggtagaaa
125220
taagctcact cacccacgca gggtactaaa gatcttacag gagcttcaac tggagcagga
125280
ggagcttttt atgcttatgt tgaatcaagt cagatacaaa aagcaattgt ccctctttgc
125340
ccaagccttt ccaattctgt gtgtcttgtt gtgtcagtgt ccacttgtgt atccttctgc
125400
aggaagaccc gccaaataga agagatggga caaaaatagg aatggtgtgt gacgacaaag
125460
ggctactgga agaacaaaag ggatacaggc cttcttgatt atctttggct ttgtacctga
125520
ggcaggagag aagagatgtc caaccagtga gatctttaag agaaaagttt gtattttaaa
125580
tgtcaatgtg cctgagaaat gtcagcttca ccacgctctt gcttcctaat gctctataca
125640
aagagggctg actatatttc ttgaagtggt gtaaaaactt agagatttta taagagaacc
125700
aggggctccc ttcacctctc ctggtccctc aggtcacata tgaaagcatt tttacaagat
125760
aggaactgga attcctcatt tctcccatgt tcctgcttgt tcttaaactt catgaagcta
125820
tttttccagc ctatggggta gttcttgctc cagtaagagg aatcttagtt gtcataatcc
125880
cttggagcct gggtttttgg agaaagagat ctccgtgccc tacagacctt ttctcaacga
125940
atgtgggaag gacctggctt taaaacacgc acacaaacac acaaataaac agacataaga
126000
tgtcatcacg aaactgccca cggatcttta ggctttctgc attgacataa atacattttc
126060
taaggggggg ggggaagaaa ttaaaaaaca cctgttaatt ttaaacacat tttttaagaa
126120
aaaaataatt aaaaaagaaa cagtgctcat gtcataagct atgttgacag ttgccagtgg
126180
aaatgttggg ttggttcaaa aaaaaaataa aagctatact atatctctct acatacagct
126240
tgcttctacc tgtgtttctt cagtgaaagg tccagggggc cactgtgggc ttcttgtgag
126300
gagacgtgac tcaggtgaag gtgtcacctc ctctcacact caggtgccaa tgtgtcagac
126360
ccagtatatt ctaagcaaaa atacttcagg aaaatgccac ttgtcaaaac ctggactttg
126420
126420
cgaagttgga agatgtaagt agtagtaaaa gctgtggtaa ttatggagga aggaggtttc
126480

tgtatcagaa aggcattggc cgtgacagac tc
126512

<210> 4
<211> 927
<212> PRT
<213> Rat

<400> 4
Met Ala Trp Leu Arg Leu Gln Pro Leu Thr Ser Ala Phe Leu His Phe
1   5   10   15
Gly Leu Val Thr Phe Val Leu Phe Leu Asn Gly Leu Arg Ala Glu Ala
            20   25   30
Gly Asp Leu Arg Asp Val Pro Ser Ala Gly Gln Asn Asn Glu Ser Cys
        35   40   45
Ser Gly Ser Ser Asp Cys Lys Glu Gly Val Ile Leu Pro Ile Trp Tyr
    50   55   60
Pro Glu Asn Pro Ser Leu Gly Asp Lys Ile Ala Arg Val Ile Val Tyr
65   70   75   80
Phe Val Ala Leu Ile Tyr Met Phe Leu Gly Val Ser Ile Ile Ala Asp
                85   90   95
Arg Phe Met Ala Ser Ile Glu Val Ile Thr Ser Gln Glu Arg Glu Val
            100   105   110
Thr Ile Lys Lys Pro Asn Gly Glu Thr Ser Thr Thr Thr Ile Arg Val
        115   120   125
Trp Asn Glu Thr Val Ser Asn Leu Thr Leu Met Ala Leu Gly Ser Ser
    130   135   140
Ala Pro Glu Ile Leu Leu Ser Leu Ile Glu Val Cys Gly His Gly Phe
145   150   155   160
Ile Ala Gly Asp Leu Gly Pro Ser Thr Ile Val Gly Ser Ala Ala Phe
                165   170   175
Asn Met Phe Ile Ile Ile Gly Ile Cys Val Tyr Val Ile Pro Asp Gly
            180   185   190
Glu Thr Arg Lys Ile Lys His Leu Arg Val Phe Phe Val Thr Ala Ala
        195   200   205
Trp Ser Val Phe Ala Tyr Ile Trp Leu Tyr Met Ile Leu Ala Val Phe
    210   215   220
Ser Pro Gly Val Val Gln Val Trp Glu Gly Leu Leu Thr Leu Phe Phe
225   230   235   240
Phe Pro Val Cys Val Leu Leu Ala Trp Val Ala Asp Lys Arg Leu Leu
                245   250   255
Phe Tyr Lys Tyr Met His Lys Arg Tyr Arg Thr Asp Lys His Arg Gly
            260   265   270
Ile Ile Ile Glu Thr Glu Gly Glu His Pro Lys Gly Ile Glu Met Asp
        275   280   285
Gly Lys Met Met Asn Ser His Phe Leu Asp Gly Asn Leu Ile Pro Leu
    290   295   300
Glu Gly Lys Glu Val Asp Glu Ser Arg Arg Glu Met Ile Arg Ile Leu
305   310   315   320
Lys Asp Leu Lys Gln Lys His Pro Glu Lys Asp Leu Asp Gln Leu Val
                325   330   335
Glu Met Ala Asn Tyr Tyr Ala Leu Ser His Gln Gln Lys Ser Arg Ala
            340   345   350
Phe Tyr Arg Ile Gln Ala Thr Arg Met Met Thr Gly Ala Gly Asn Ile
        355   360   365
Leu Lys Lys His Ala Ala Glu Gln Ala Lys Lys Thr Ala Ser Met Ser
    370   375   380
Glu Val His Thr Asp Glu Pro Glu Asp Phe Ala Ser Lys Val Phe Phe
385   390   395   400